

DOCUMENT RESUME

ED 411 321

TM 027 390

TITLE Measuring the Quality of Schools = Mesurer la Qualite des Etablissements Scolaires.

INSTITUTION Organisation for Economic Cooperation and Development, Paris (France). Centre for Educational Research and Innovation.

ISBN ISBN-92-64-04355-1

PUB DATE 1995-00-00

NOTE 238p.

AVAILABLE FROM OECD Washington Center, 2001 L Street, N.W., Suite 650, Washington, DC 20036-4922; phone: 1-800-456-6323 (\$29, plus \$4.50 shipping and handling).

PUB TYPE Collected Works - General (020) -- Multilingual/Bilingual Materials (171)

LANGUAGE English, French

EDRS PRICE MF01/PC10 Plus Postage.

DESCRIPTORS *Academic Achievement; *Accountability; Comparative Analysis; *Educational Quality; *Elementary Secondary Education; Equal Education; Foreign Countries; Formative Evaluation; Higher Education; *International Education; *Measurement Techniques; Performance Factors

IDENTIFIERS Educational Indicators; *Organisation for Economic Cooperation Development

ABSTRACT

As part of the Organisation for Economic Cooperation and Development project on international educational indicators, four networks were formed to represent different domains. This collection contains background papers that were prepared to deal with some of the fundamental questions about process indicators examined by Network "C," the network that, under the leadership of The Netherlands, was given the task of measuring indicators of schools and school processes. The following papers are included: (1) "Internationally Comparable Indicators of Educational Programmes and Processes: Identification, Measurement and Interpretation" (Jaap Scheerens); (2) "Comparer les structure de decision des systemes educatifs: un bilan de l'approche quantitative (Comparison of Decision-Making Structures in Education Systems: A Review of the Quantitative Approach)" (Denis Meuet, Jacques Prod'hom, and Eugen Stocker); (3) "The Selection and Definition of International Indicators on Teachers" (Jaap Scheerens); (4) "Curriculum Indicators in International Comparative Research" (Willem J. Pelgrum, Joke Voogt, and Tjeerd Plomp); (5) "Process Indicators on School Functioning and the Generalisability of School Factor Models across Countries" (Bert Creemers); (6) "The Development of Indicators on Equity in Education" (Alan Gibson and Denis Meuret); (7) "The Search for Indicators on the Effective Functioning of Tertiary Education Institutions" (Alan Gibson); (8) "The Use of Performance Indicators in School Improvement and Accountability" (Carmo Climaco); (9) "Perspectives on Decision-Making in Education Systems and the Development of the Network C Locus of Decision-Making Survey" (Darrel Drury, Laura Hersch Salganik, and Marilyn M. McMillen); (10) "La mise en place d'un dispositif d'indicateurs pour le pilotage des etablissements secondaires francais (Implementation of a System of Indicators for the Steering of French Secondary Schools)" (Jean-Calude Emin); and (11) "Interpretation and Use of Education Indicators" (Erik

+++++ ED411321 Has Multi-page SFR---Level=1 +++++

Wallin). Each chapter contains references. (Contains 2 boxes, 17 figures, and
19 tables.) (SLD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

OECD • CODE

ED 411 321

TM027390

TM

CENTRE FOR EDUCATIONAL RESEARCH AND INNOVATION
CENTRE POUR LA RECHERCHE ET L'INNOVATION
DANS L'ENSEIGNEMENT

INDICATORS OF EDUCATION SYSTEMS
INDICATEURS DES SYSTÈMES D'ENSEIGNEMENT

MEASURING THE QUALITY OF SCHOOLS



MESURER LA QUALITÉ DES ÉTABLISSEMENTS SCOLAIRES

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
IN OTHER THAN PAPER COPY
HAS BEEN GRANTED BY

J. R. WEST

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

* This document has been reproduced as
received from the person or organization
originating it.

☐ Minor changes have been made to
improve reproduction quality.

• Points of view or opinions stated in this
document do not necessarily represent
official OERI position or policy.

2

BEST COPY AVAILABLE

INDICATORS OF EDUCATION SYSTEMS

**INDICATEURS DES SYSTÈMES
D'ENSEIGNEMENT**

**MEASURING
THE QUALITY
OF SCHOOLS**



**MESURER
LA QUALITÉ
DES ÉTABLISSEMENTS
SCOLAIRES**

**ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT
ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUES**

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

ORGANISATION DE COOPÉRATION ET DE DÉVELOPPEMENT ÉCONOMIQUES

Pursuant to Article 1 of the Convention signed in Paris on 14th December 1960, and which came into force on 30th September 1961, the Organisation for Economic Co-operation and Development (OECD) shall promote policies designed:

- to achieve the highest sustainable economic growth and employment and a rising standard of living in Member countries, while maintaining financial stability, and thus to contribute to the development of the world economy;
- to contribute to sound economic expansion in Member as well as non-member countries in the process of economic development; and
- to contribute to the expansion of world trade on a multilateral, non-discriminatory basis in accordance with international obligations.

The original Member countries of the OECD are Austria, Belgium, Canada, Denmark, France, Germany, Greece, Iceland, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The following countries became Members subsequently through accession at the dates indicated hereafter: Japan (28th April 1964), Finland (28th January 1969), Australia (7th June 1971), New Zealand (29th May 1973) and Mexico (18th May 1994). The Commission of the European Communities takes part in the work of the OECD (Article 13 of the OECD Convention).

En vertu de l'article 1^{er} de la Convention signée le 14 décembre 1960, à Paris, et entrée en vigueur le 30 septembre 1961, l'Organisation de Coopération et de Développement Économiques (OCDE) a pour objectif de promouvoir des politiques visant :

- à réaliser la plus forte expansion de l'économie et de l'emploi et une progression du niveau de vie dans les pays Membres, tout en maintenant la stabilité financière, et à contribuer ainsi au développement de l'économie mondiale ;
- à contribuer à une saine expansion économique dans les pays Membres, ainsi que les pays non membres, en voie de développement économique ;
- à contribuer à l'expansion du commerce mondial sur une base multilatérale et non discriminatoire conformément aux obligations internationales.

Les pays Membres originaires de l'OCDE sont : l'Allemagne, l'Autriche, la Belgique, le Canada, le Danemark, l'Espagne, les États-Unis, la France, la Grèce, l'Irlande, l'Islande, l'Italie, le Luxembourg, la Norvège, les Pays-Bas, le Portugal, le Royaume-Uni, la Suède, la Suisse et la Turquie. Les pays suivants sont ultérieurement devenus Membres par adhésion aux dates indiquées ci-après : le Japon (28 avril 1964), la Finlande (28 janvier 1969), l'Australie (7 juin 1971), la Nouvelle-Zélande (29 mai 1973) et le Mexique (18 mai 1994). La Commission des Communautés européennes participe aux travaux de l'OCDE (article 13 de la Convention de l'OCDE).

© OECD 1995
OCDE

Applications for permission to reproduce or translate all or part of this publication should be made to:

Les demandes de reproduction ou de traduction totales ou partielles de cette publication doivent être adressées à :

Head of Publications Service, OECD

M. le Chef du Service des Publications, OCDE

2, rue André-Pascal, 75775 PARIS CEDEX 16, France.

Foreword

The General Assembly of the OECD project on International Indicators of Education Systems (INES) met in Lugano, Switzerland, in September 1991. That assembly reviewed the first set of OECD education indicators, and recommended that they be published in *Education at a Glance : OECD Indicators*. After the decision by the CERI Governing Board and the Education Committee in early 1992 to continue to develop education indicators, four networks with voluntary country participation were invited to pursue the conceptual and methodological work needed for the measurement of new indicators in different domains:

- Network A, led by the United States, took up the challenge of developing and measuring indicators of student learning outcomes;
- Network B, with strong support from Sweden, was requested to develop measures of education and labour market destinations;
- Network C, under the leadership of the Netherlands, was given the task of measuring indicators of schools and school processes;
- Network D, supported by the United Kingdom, was invited to chart the expectations and attitudes to education of the various stakeholder groups in OECD societies.

The General Assembly will convene again in June 1995 when Member countries and the people involved in the INES project will take stock of what has been achieved, examine the organisational framework of the sets of indicators produced so far, and explore the possibilities for further developments. To facilitate this review, each network has prepared a report describing the conceptual, methodological, and policy problems encountered in constructing their respective clusters of indicators. The four reports are presented to the INES General Assembly as background and reference documents. Collectively, they offer a rich account of innovations, successes and failures in indicator development at the OECD, information that is essential for a review of perspectives and options for possible future work.

The content of each report has been discussed and endorsed during plenary sessions of the networks. The contributors are mostly members and national delegates to the networks, although in some cases distinguished independent experts have also contributed. For each network, members reviewed all the papers and suggested modifications where needed.

The preparation of this report was made possible thanks to the support received from the Netherlands, the lead country of Network C, particularly the Netherlands Ministry of Education and Science, which contributed generously and co-ordinated the production of this report. This volume was prepared by the network Chair, Jaap Scheerens of the University of Twente, the Netherlands, in co-operation with Norberto Bottani and Albert Tuijnman of the OECD Secretariat.

This report is published on the responsibility of the Secretary-General of the OECD. It represents the views of the authors and does not necessarily reflect those of the OECD or of its Member countries.

Avant-propos

L'Assemblée générale du projet de l'OCDE sur les indicateurs internationaux des systèmes d'enseignement (INES) s'est réunie à Lugano, en Suisse, en septembre 1991. Elle a passé en revue la première série d'indicateurs de l'enseignement de l'OCDE et recommandé qu'ils soient publiés dans *Regards sur l'éducation : les indicateurs de l'OCDE*. Après que le Comité directeur du CERI et le Comité de l'éducation eurent décidé au début de 1992 de poursuivre l'élaboration d'indicateurs de l'enseignement, quatre réseaux, auxquels les pays participent volontairement, ont été invités à effectuer les travaux conceptuels et méthodologiques nécessaires pour mesurer de nouveaux indicateurs dans différents domaines :

- le Réseau A, mené par les États-Unis, s'est donné pour mission d'élaborer et de mesurer des indicateurs des acquis des élèves et étudiants;
- le Réseau B, qui bénéficie d'une importante aide de la Suède, a été invité à mettre au point des mesures relatives à l'enseignement et aux débouchés professionnels;
- le Réseau C, dirigé par les Pays-Bas, s'est vu confier la tâche de mesurer les indicateurs des établissements et des processus scolaires;
- le Réseau D, soutenu par le Royaume-Uni, a été invité à décrire les attentes et les attitudes à l'égard de l'enseignement des diverses parties prenantes dans les pays de l'OCDE.

L'Assemblée générale se réunira de nouveau en juin 1995, ce qui permettra aux pays Membres et aux personnes participant au projet INES de faire le bilan de ce qui a été accompli, d'examiner l'organisation des séries d'indicateurs élaborés jusqu'à présent, et d'étudier les prolongements possibles. Pour faciliter cet examen, chaque réseau décrit dans un rapport les problèmes conceptuels, méthodologiques et politiques auxquels il s'est heurté dans la construction de sa propre série d'indicateurs. Les quatre rapports sont présentés à l'Assemblée générale du projet INES à titre de documents de référence. Ensemble, ils rendent compte en détail des innovations introduites avec ou sans succès dans la construction d'indicateurs à l'OCDE, autant d'informations qui sont essentielles pour étudier les perspectives et les options qui s'offrent en vue de la poursuite des travaux.

Le contenu de chaque rapport a été examiné et approuvé au cours des sessions plénières des réseaux. Les auteurs sont dans leur majorité des membres des réseaux et des personnes déléguées près d'eux par les pays, encore que dans certains cas des experts

indépendants de haut niveau aient aussi apporté leur concours. Pour chacun des réseaux, les membres ont passé en revue tous les rapports et, le cas échéant, suggéré des modifications.

Ce rapport a pu être élaboré grâce à l'aide reçue des Pays-Bas, qui assurent la direction du Réseau C, et en particulier du ministère néerlandais de l'Éducation et de la Science, qui a généreusement contribué à sa publication et en a assuré la coordination. Cet ouvrage a été établi par le président du réseau, Jaap Scheerens, professeur à l'Université de Twente, aux Pays-Bas, en coopération avec Norberto Bottani et Albert Tuijnman du Secrétariat de l'OCDE.

Ce rapport reflète les opinions des auteurs et ne représente pas nécessairement le point de vue de l'OCDE ni celui de ses pays Membres. Il est publié sous la responsabilité du Secrétaire général de l'OCDE.

Table of Contents/Table des matières

Introduction by Jaap Scheerens	13
--------------------------------------	----

Chapter/Chapitre 1

Internationally Comparable Indicators of Educational Programmes and Processes: Identification, Measurement and Interpretation

Indicateurs comparables au plan international sur les programmes et processus de l'enseignement : identification, mesure et interprétation

Jaap Scheerens

1. Introduction	19
2. Background Information	19
3. Identification of Process Indicators	20
The Human Relations Model	23
The Open System Model	24
The Internal Process Model	25
4. Data Collection Strategies	26
5. Interpretation and Use of Process Indicators for Policy-Making	28
6. Conclusion	30
References	31

Chapter/Chapitre 2

Comparer les structures de décision des systèmes éducatifs : un bilan de l'approche quantitative

Comparison of Decision-Making Structures in Education Systems: A Review of the Quantitative Approach

Denis Meuret, Jacques Prod'homme et Eugen Stocker

1. Introduction	35
2. Autres approches de l'organisation des systèmes éducatifs	36
Un environnement politique et intellectuel différent	37
Des approches différentes	40
Les résultats	43
Une source de suggestions	45

3. Comparaison avec d'autres caractéristiques des systèmes éducatifs	48
La réussite à la fin du second cycle	48
Les dépenses par élève	49
Le niveau de lecture des élèves de 14 ans	51
4. Conclusion	52
Références	53
Annexe : Relations entre le niveau de lecture à 14 ans et le poids des différents niveaux administratifs dans la structure de décision	55

Chapter/Chapitre 3

The Selection and Definition of International Indicators on Teachers **Sélection et définition des indicateurs internationaux sur les enseignants**

Jaap Scheerens

1. Introduction	61
2. Rationales for Considering Teacher Indicators	62
Issues that Presently Confront the Teaching Profession	62
Macro-level Description	64
Effectiveness and Efficiency	65
The Political Economy of the Teaching Profession: System-Level Conditions that Affect Teachers' Motivation	66
3. Methods Used in the OECD INES Survey on Teachers	69
Standard Definitions	69
Procedure	69
4. Discussion of Results	69
Teaching Time	69
Duration of Teachers' Education	71
Teacher Compensation	72
Teacher Characteristics (Age and Sex)	73
5. Conclusion	76
References	78

Chapter/Chapitre 4

Curriculum Indicators in International Comparative Research **Indicateurs relatifs aux programmes scolaires dans la recherche comparative internationale**

Willem J. Pelgrum, Joke Voogt and Tjeerd Plomp

1. Introduction	83
2. A Theoretical Perspective	84
3. The Empirical Scene	87
Content	88
Time	88
Participants	89

Instructional Characteristics	89
4. Reliability and Validity of Curriculum Indicators	90
5. Summary, Conclusions and Recommendations	92
References	95
Annex: Overview of indicators and sources	98

Chapter/Chapitre 5

Process Indicators on School Functioning and the Generalisability of School Factor Models Across Countries

Indicateurs de processus sur le fonctionnement de l'école et la possibilité d'application transnationale des modèles de facteurs scolaires

Bert Creemers

1. Introduction: The Value of Process Indicators	105
2. Background	106
3. The International School Effectiveness Research Programme (ISERP)	106
4. A Study of "Archetypal" Schools in a Range of Countries	107
5. Research Questions and Design of the Study	110
6. First Results	111
7. Conclusion	116
References	118

Chapter/Chapitre 6

The Development of Indicators on Equity in Education *La mise au point d'indicateurs sur l'équité dans l'enseignement*

Alan Gibson and Denis Meuret

1. Introduction	122
2. The Concern for Equity	123
3. Equity and Equality	125
4. The Development of Indicators on Equity	127
References	128
Annex 6.1: Variables useful for calculating indicators on equity	129
Annex 6.2: Indicators in Education at a Glance (OCDE, 1993) showing disparity by gender, socio-economic group, race, religion or other categories listed in Annex 6.1	131

Chapter/Chapitre 7

The Search for Indicators on the Effective Functioning of Tertiary Education Institutions
*La recherche d'indicateurs du fonctionnement efficace des établissements
d'enseignement supérieur*

Alan Gibson

1. Introduction	134
2. The Value of Indicators on the Functioning of Tertiary Education Institutions	135
3. Existing Work on Indicators on the Functioning of Tertiary Education Institutions	136
4. The Nature of Indicators Needed in Future	138
References	140
Annex: Examples of possible indicators of effective functioning in tertiary education	141

Chapter/Chapitre 8

The Use of Performance Indicators in School Improvement and in Accountability
Les indicateurs de performance au service de l'amélioration et de la transparence de l'école

Carmo Clímaco

1. Introduction	146
The Quality Debate	147
Initiatives for the Monitoring of Quality	149
The Organisation of School Information Systems	149
2. Method	151
The Initial Model of Indicators	152
The Selection of Schools	153
Results of the Initial Survey	154
The Development of Qualitative Indicators	156
Revision of the Set of Indicators	156
Accountability Indicators	157
3. Some Results	158
The Context Indicators	159
The Results Indicators	160
The Factors Related to Teachers	162
Teachers as Human Resources	163
Participation in School Life by Different Groups	163
4. From Improvement to Accountability	169
Development of a Common Core of Indicators	170
Testing of the Common Core of Indicators	170
5. Conclusion	173
References	175

Chapter/Chapitre 9

Perspectives on Decision-Making in Education Systems and the Development of the Network C Locus of Decision-Making Survey

Considérations sur la prise de décision dans les systèmes d'enseignement et sur l'évolution de l'enquête menée par le Réseau C sur les niveaux où se prennent les décisions

Darrel Drury, Laura Hersch Salganik and Marilyn M. McMillen

1. Introduction	182
2. Theoretical Perspectives	183
Institutional Theory	183
Educational Organisations as Political Systems	186
Synthesis of the Perspectives	188
3. The Network C Locus of Decision-Making Survey	190
The Current Survey	191
Extending the Network C Survey	193
4. Conclusion	195
References	197

Chapter/Chapitre 10

La mise en place d'un dispositif d'indicateurs pour le pilotage des établissements secondaires français

Implementation of A System of Indicators for the Steering of French Secondary Schools

Jean-Claude Emin

1. Le contexte	203
Une nécessité légale	203
Une demande sociale de plus en plus exigeante et pressante quant aux résultats et à l'efficacité des établissements	204
Les premières réalisations ont permis de développer une culture de l'évaluation au niveau des établissements secondaires	205
Le développement de pratiques d'évaluation chez les enseignants à partir d'outils standardisés	206
L'auto-évaluation des établissements reste difficile et très minoritaire	207
2. Les orientations retenues pour la mise en place du dispositif	207
Se placer dans une optique d'aide et de service rendu aux établissements	208
Un outil qui permet à l'établissement, à la fois, de se rendre compte de son fonctionnement et d'en rendre compte	208
Un outil qui permet une évaluation qui se fonde, à la fois, sur une base d'indicateurs relativement complète et commune à l'ensemble des établissements et sur des indicateurs spécifiques à chaque établissement et à son projet	209
Une généralisation rapide et un important soutien en formation	210
3. La batterie d'indicateurs standard	210

Décrire, aussi complètement que possible, tous les aspects d'un modèle de fonctionnement des établissements scolaires	211
Dépasser l'appréciation des résultats des établissements pour évaluer leur « valeur ajoutée »	211
Donner aux établissements des moyens de comprendre leur fonctionnement, d'apprécier leurs contraintes et leurs ressources	221
Un indicateur de ressources : les moyens horaires des établissements	222
La proportion d'heures d'enseignement dispensées à un élève	223
L'importance des services liés à la vie scolaire	226
4. Les premières réactions à la mise en place du dispositif et les perspectives	227
Premières réactions positives en ce qui concerne l'outil d'auto-évaluation et de pilotage et quelques états d'âme par rapport à l'évaluation externe	227
Des perspectives d'évolution de plusieurs ordres	228
Notes	230
Références	231

Chapter/Chapitre 11

Interpretation and Use of Education Indicators *Interprétation et exploitation des indicateurs de l'enseignement*

Erik Wallin

1. Introduction	234
2. Network C – A Context for Interpretation	235
3. INES – Its Context for Use of Data	237
Restructuring of the Education System	237
Decision-Makers and Teachers as Decision-Makers	238
The Encounter between Knowledge and User	238
4. Back to INES Data	239
References	240

Introduction

by

Jaap Scheerens

Department of Education, University of Twente, the Netherlands

Indicators on the functioning of schools and school systems are the domain of Network C of the OECD CERI project on International Indicators of Education Systems, the INES project. Included in this category are indicators about human resources in education, decision-making processes, the curriculum, and organisational and instructional processes at school level.

The place of this category of indicators within a larger framework of a *system* of indicators, is to some extent on the input side (*i.e.* input of human resources in education), but it is mainly associated with "throughput" or "process" – in other words, with the process of transforming inputs into outputs. The general term "process indicators" is therefore frequently used to describe the domain covered by Network C.

The network structure of the INES project unites representatives from OECD countries in groups that guide the process of developing indicators as well as the actual information gathering and computation of indicators. In most cases, the network leaders have a small supporting academic staff to assist in the construction of formats and instruments, and the co-ordination of data collection and analysis of data. In the case of Network C, the leading team was situated first at the *Département de l'Évaluation et de la Prospective* of the French Ministry of Education and later, from 1992, at the Centre of Applied Research in Education (OCTO) of the University of Twente in the Netherlands.

Network C started in 1989. The first phases of the work were devoted to clarifying the nature of "process" indicators and to the selection and conceptualisation of indicators. The first two indicators that were actually calculated were teacher : pupil ratio and the proportion of the labour force in a country which is actively engaged in education. The definition and data collection for indicators on decision-making distributed over administrative levels was carried out by a Franco-Swiss sub-group of the network.

During the present phase of the INES project, data on teachers and the intended curriculum were collected, and data gathered for the Reading Literacy Study by the International Association for the Evaluation of Educational Achievement (IEA) were used to compute indicators on school processes. Presently, a pilot study for a survey of

schools is being conducted to improve the international database for indicators on school processes. The network is also exploring new areas such as indicators on equity, and process indicators concerning the functioning of higher education institutes.

This book consists of a series of background papers that deal with some of the more fundamental questions regarding process indicators. What is the rational basis for selecting and identifying indicators? What is the political relevance? Is there an empirical basis that supports the significance of selected indicators? What are the strengths and limitations of the current set of indicators? What developments are relevant for the future?

The contents of the book cover five main areas:

- the identification and measurement of process indicators on the functioning of schools;
- rationales and measurement strategies concerning indicators on teachers and the curriculum;
- the measurement and interpretation of indicators on decision-making across administrative levels;
- new areas for development of process indicators: equity and tertiary education;
- the use of indicator systems at the level of national education systems.

The final chapter is a reflection on the place of the indicators developed in Network C within the framework of the total INES indicator set.

In dealing with these areas, we cover important general issues of education such as the empirical status of causal models of educational effectiveness, various interpretations of organisational effectiveness when applied to educational institutes, the strategic importance of the human resources factor in education, theoretical models of the functioning of schools as organisations, various conceptualisations of the curriculum, issues of centralisation and decentralisation, perspectives on the equity issue in education and methodological questions concerning the collection and analysis of data for process indicators.

Of course, the substance of the book is to be judged by its readers; as far as the process of its creation is concerned, it is a successful example of good co-operation within the network.

Introduction

par

Jaap Scheerens

Département de l'éducation, Université de Twente, Pays-Bas

Les indicateurs du fonctionnement de l'école et du système scolaire relèvent du Réseau C du projet OCDE/CERI sur les indicateurs internationaux des systèmes d'enseignement (projet INES). Sont inclus dans cette catégorie les indicateurs des ressources humaines consacrées à l'enseignement, des processus de décision, des programmes et des processus d'organisation et d'enseignement au niveau de l'école.

Cette catégorie d'indicateurs qui s'insèrent dans le cadre plus large d'un système d'indicateurs concerne, dans une certaine mesure, l'apport de ressources (autrement dit les ressources humaines consacrées à l'éducation), mais a trait pour l'essentiel au « volume d'activité » ou au « processus » – en d'autres termes à la transformation des ressources en résultats. Le terme générique « d'indicateurs de processus » est donc souvent utilisé pour décrire le domaine d'activité du Réseau C.

La structure en réseaux du projet INES rassemble des représentants des pays de l'OCDE au sein de groupes qui président à l'élaboration d'indicateurs ainsi qu'à la collecte effective des données et au calcul des indicateurs. Dans la plupart des cas, les responsables de réseau disposent du concours d'un petit nombre d'universitaires qui aident à établir des schémas et des instruments, à coordonner la collecte des données et à les analyser. Dans le cas du Réseau C, l'équipe dirigeante a opéré d'abord au Département de l'évaluation et de la prospective du ministère français de l'Éducation nationale, puis, à partir de 1992, au Centre de recherche appliquée à l'éducation (OCTO) de l'Université de Twente aux Pays-Bas.

Le Réseau C est entré en activité en 1989. Les premières phases de ses travaux ont consisté à définir la nature des indicateurs de « processus », à sélectionner et conceptualiser les indicateurs. Les deux premiers indicateurs qui ont été effectivement construits étaient le taux d'encadrement et la proportion d'actifs d'un pays engagés dans des activités éducatives. C'est un sous-groupe franco-suisse du réseau qui s'est chargé des définitions et de la collecte des données pour construire les indicateurs de processus de décision aux divers niveaux de l'administration.

Durant la phase du projet INES actuellement en cours, on a recueilli des données sur les enseignants et sur les programmes envisagés, et on a utilisé des données émanant de l'étude sur le niveau de lecture réalisée par l'International Association for the Evaluation of Educational Achievement (IEA) pour établir des indicateurs de processus concernant l'école. Une étude pilote en vue d'une enquête auprès des écoles est en cours afin d'améliorer la base internationale de données nécessaire à l'élaboration d'indicateurs des processus scolaires. Le réseau explore aussi de nouveaux domaines comme les indicateurs d'équité et les indicateurs de processus concernant le fonctionnement des établissements d'enseignement supérieur.

Cet ouvrage rassemble une série de documents de base qui traitent de certaines des questions essentielles liées aux indicateurs de processus. Sur quelle base rationnelle se fonde-t-on pour choisir et définir des indicateurs? Quel en est l'intérêt politique? L'importance de certains indicateurs est-elle empiriquement justifiée? Quels sont les avantages et les inconvénients de la série actuelle d'indicateurs? Quelles améliorations seraient nécessaires pour l'avenir?

Cinq grands domaines sont étudiés :

- la définition et le calcul des indicateurs de processus concernant le fonctionnement de l'école;
- la raison d'être et les méthodes de calcul des indicateurs sur les enseignants et les programmes;
- le calcul et l'interprétation des indicateurs de processus de décision aux différents niveaux de l'administration;
- les domaines nouveaux susceptibles de faire l'objet d'indicateurs de processus : l'équité et l'enseignement supérieur;
- l'utilisation de systèmes d'indicateurs dans les systèmes d'enseignement nationaux.

Le dernier chapitre analyse la place qu'occupent les indicateurs mis au point par le Réseau C dans la série d'indicateurs du projet INES.

En abordant ces domaines, nous évoquons d'importantes questions générales relatives à l'enseignement, notamment la base empirique des modèles de causalité touchant l'efficacité de l'enseignement, les diverses interprétations de l'efficacité d'organisation appliquée aux établissements d'enseignement, l'importance stratégique des ressources humaines dans l'enseignement, les modèles théoriques de fonctionnement des écoles en tant qu'organisations, les divers concepts de programme, les problèmes de centralisation et de décentralisation, les perspectives d'équité dans l'enseignement et les problèmes méthodologiques que posent la collecte et l'analyse de données utilisées pour la construction d'indicateurs de processus.

Il appartient naturellement au lecteur de juger de la teneur de cet ouvrage; sa rédaction, quant à elle, constitue un exemple de coopération réussie au sein du réseau.

Chapter/Chapitre 1

**Internationally Comparable Indicators of Educational Programmes
and Processes: Identification, Measurement and Interpretation**
*Indicateurs comparables au plan international sur les programmes
et processus de l'enseignement : identification, mesure
et interprétation*

by

Jaap Scheerens

Department of Education, University of Twente, the Netherlands

Process indicators of school functioning form a category of education indicators that differ from the more common indicators dealing with global, low-inference statistics defined at a high level of aggregation, which are intended for use by decision-makers at the national level. Within the framework of the OECD project on International Indicators of Education Systems (INES), the possibilities for the development of indicators on school inputs and processes are being explored. This chapter points out some of the problems in identifying, measuring, interpreting and defining the use of process indicators on school functioning for policy-making. Among the solutions offered to overcome these problems are 1) the use of school effectiveness modelling and research as an important resource for selecting process indicators, 2) secondary analyses, or collaboration in current international surveys, in order to obtain basic data, 3) international panels of experts to determine the standards according to which process indicators should be interpreted, and 4) a differentiated approach to the use of process indicators for policy-making, since not all process indicators will be relevant at all policy-making levels.

Experience so far has shown that there are many difficulties in the use of data available from international surveys, although there are also unique advantages in employing available data sets. Network C has therefore been investing considerable energy in the preparation and conduct of specially designed data-collection strategies, such as surveys of schools, while at the same time keeping the use of available international data on the agenda.

*
* *
*

Note de synthèse

Les indicateurs de processus du fonctionnement de l'école constituent une catégorie d'indicateurs de l'enseignement qui s'écarte de l'idée généralement admise selon laquelle les indicateurs reposent sur des statistiques globales, à faible degré d'inférence, définies à un niveau élevé d'agrégation, et destinées à servir aux décideurs à l'échelon national. Dans le cadre du projet de l'OCDE sur les indicateurs internationaux des systèmes d'enseignement (projet INES), on explore les possibilités d'élaborer des indicateurs sur les moyens et les processus éducatifs. On énumère dans ce chapitre quelques-uns des problèmes que soulèvent l'identification, la mesure et l'interprétation des indicateurs de processus du fonctionnement de l'école, ainsi que la définition de l'usage qui peut en être fait par les pouvoirs publics. Parmi les solutions proposées figurent:

- le recours à la modélisation et à la recherche sur l'efficacité de l'école qui constituerait un important moyen de sélection des indicateurs de processus, en partant d'un cadre plus général, inspiré par la documentation consacrée à l'efficacité de l'entreprise, modèle implicite de la plupart des recherches sur l'efficacité de l'école;
- les analyses secondaires ou le rattachement aux enquêtes actuellement menées au plan international pour obtenir des données de base;
- la création de groupes d'experts internationaux pour définir les normes servant à interpréter les indicateurs de processus;
- une approche nuancée de l'utilisation des indicateurs de processus en vue de l'élaboration de l'action des pouvoirs publics (dans la mesure où les indicateurs de processus ne sont pas utilisables à tous les niveaux de décision, et en tenant compte des différences entre pays dans ce domaine).

On se préoccupe en premier lieu de trouver des stratégies de collecte des données qui soient à la fois envisageables et capables de donner lieu à des indicateurs sur les programmes et processus de l'enseignement qui se prêtent aux comparaisons internationales. Il ressort de l'expérience acquise jusqu'à présent que l'utilisation des données fournies par les enquêtes internationales soulève de nombreuses difficultés bien que le recours à des séries de données déjà disponibles présente d'exceptionnels avantages. C'est pourquoi le Réseau C a consacré beaucoup d'énergie à la préparation et à la conduite de procédures spécialement conçues pour la collecte de données, telles que les enquêtes sur les établissements scolaires, tout en gardant à l'ordre du jour l'exploitation des séries de données internationales disponibles.

*
* *
*

1. Introduction

Education indicators are statistics that are useful for planning, management and policy-making (OECD, 1991, p. 19). Usually, indicators are thought of as simple, global and lean, and defined at a high level of aggregation. The main level of administrative use is policy-making and management at the national level.

This view of indicators is predominant in the OECD project on International Indicators of Education Systems (INES). Yet, in this project there is also room for some more innovative types of education indicators. One of these types is referred to as *indicators on educational programmes and processes*, in which indicators on the functioning of schools are explicitly being considered. As pointed out elsewhere (Scheerens, 1990, 1992), thinking in terms of *process indicators* of school functioning also brings the issue of indicator use by lower-level decision-makers (school districts or even schools) to the fore.

In this chapter, possibilities as well as specific problems in selecting, measuring and interpreting process indicators of school functioning will be described.

2. Background Information

Within the OECD, the Centre for Educational Research and Innovation (CERI) focuses on innovations in educational research, teaching practices, indicators, and human resources. In 1988, the CERI started a long-term project on education indicators, the INES project. The ultimate aim of the INES project is to provide inputs for an international education monitoring system, which will give an overall picture of the education system of countries.

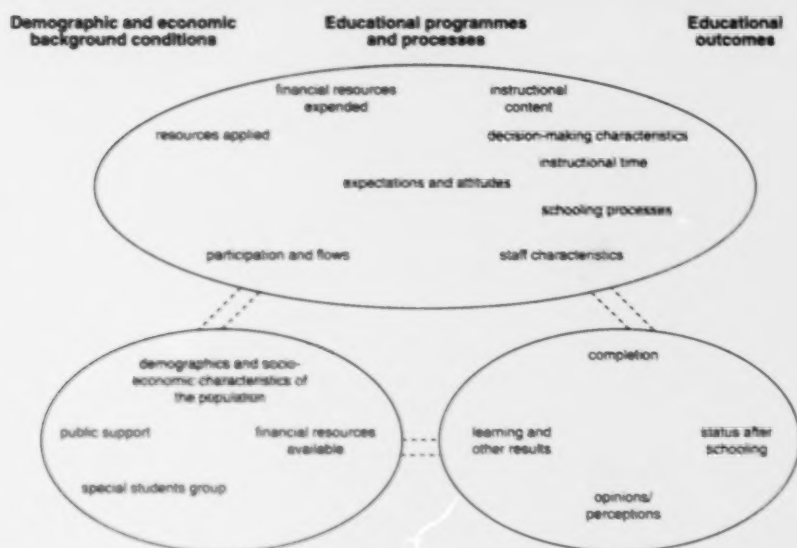
Most OECD countries (25 industrialised countries) show a growing interest in the quality of their national education systems and therefore in the activities of the INES project. They expect the project to give more insight into their own education systems, and to provide greater opportunities for comparison with the systems of other countries, than is possible on the basis of the information in current educational statistics.

The framework for an ideal set of indicators, which was developed in the exploratory phase of the INES project, is given in Figure 1.1.

Figure 1.1 should be read as a conceptual map and definitely not as a "model", because this would suggest causal connections among indicators, which really is too strong an assumption.

In making a selection from the ideal set of indicators, practical considerations such as measurement restrictions, data-collection problems and sensitivity to policy, played a relatively important role. In the first two phases of the INES project (1988-91), a consensus was reached on the selection and definition of a set of indicators. Some of the proposed indicators have actually been calculated (OECD, 1992).

Figure 1.1. "Ideal" set of OECD international indicators of education
Ensemble « idéal » d'indicateurs internationaux de l'éducation de l'OCDE



Source : OECD.

3. Identification of Process Indicators

In phases I and II of the INES project, Network C on educational programmes and processes has developed a set of twelve indicators on features of schools and school systems. They are:

- C1 teacher : pupil ratio;
- C2 between-school variation in teaching staff;
- C3 percentage of the labour force in education;
- C4 hours of instruction per student;
- C5 time on task;
- C6 topic coverage;
- C7 school leadership;
- C8 staff co-operation;
- C9 differentiated and integrated learning;
- C10 success-oriented ethos among students;
- C11 locus of decision-making (and school autonomy);
- C12 modes of decision-making.

Indicators C1, C2, C3, C11 and C12 are process indicators defined at the *system level*; the other indicators are defined at the *school or class level*.

The rationale used in identifying process indicators can be found in empirical data and models of school effectiveness. This means that *student academic achievement* in the basic skills is the central *effectiveness criterion* and organisational, curricular and instructional correlates of this criterion are selected as process indicators.

This selection of relevant process indicators proceeds from the traditional view on how indicators should be used. Indicator sets are basically meant to serve as "calculating" models to prepare (or justify) certain policy changes. Ideally, sets of indicators must then be built on an existing knowledge base of causal relationships between the context, input, process and output of the system under consideration. In the development of a set of indicators on educational programmes and processes, such as that of INES Network C, the underlying assumption is that there is a consistent knowledge base on causal relationships, from which important process indicators could be derived. Unfortunately, this is not the case: many studies in school effectiveness research in diverse countries have shown that the existing empirical base of effective school characteristics is not as solid as we would wish. In Table 1.1, an overview derived from Scheerens (1992) is given of empirical support for various educational process variables (at school and class level).

Table/Tableau 1.1.

Strength of relationship between educational process variables and outcomes
Importance du lien entre les variables du processus éducatif et les résultats

Characteristics	Multiple empirical research confirmation	A reasonable empirical basis	Doubtful empirical confirmation	Hypothetical
Structured teaching	x			
Effective learning time	x			
Opportunity to learn		x		
Pressure to achieve		x		
High expectations		x		
Pedagogical leadership			x	
Assessment			x	
School climate			x	
Staff recruitment				x
Organisational/ structural preconditions			x	
Physical/material school characteristics		o		
Descriptive context characteristics			x	
External stimuli to make schools effective				x
Parental involvement		x		

x, a meaningful influence.
o, a more marginal influence.
Source: Scheerens (1992).

From Table 1.1 it can be seen that a strong empirical basis exists for only two educational characteristics. The other relationships mentioned are judged to be moderate, weak or even hypothetical.

Not only are the relationships between these process variables and output in terms of student achievement not consistently supported by empirical research, there may also be uncertainty on the direction of causality in the existing relationships. This implies that the interpretation of figures on process indicators will cause severe problems. Later on in this chapter, the interpretation and use of process indicators for policy-making will be discussed in more detail.

A comparison of the variables listed in Table 1.1 with the school and class variables defined as indicators for Network C, shows that some indicators are included in the Network C indicator set, although there is only a weak empirical basis for these indicators to be related to student output (for example, school leadership). On the other hand, not all indicators that might be of importance according to Table 1.1 are included in the indicator set developed during the first two phases of the INES project (for example, assessment capacities of schools).

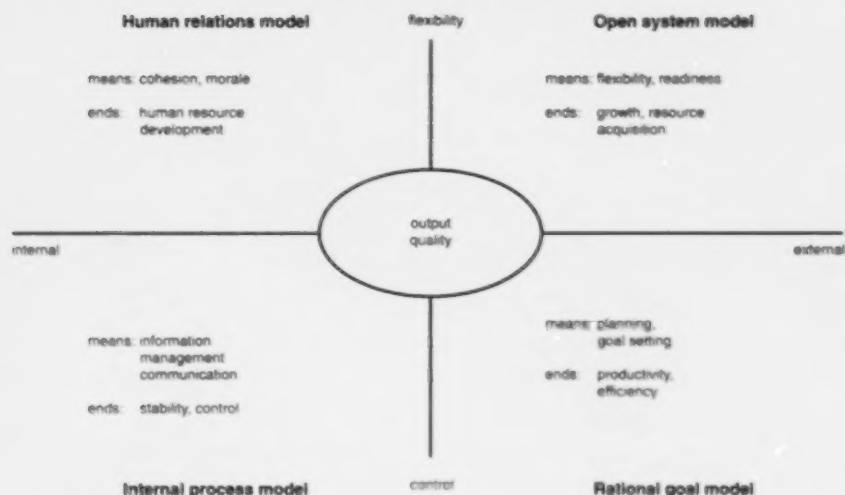
In organisational theory, models such as this school effectiveness model are seen as just one of several effectiveness perspectives. The perspective into which the school effectiveness model fits is referred to as the *rational goal model*, in which productivity and efficiency are the central criteria.

Alternative models are the *human relations model*, with human resource development as a central criterion; the *open system model*, with growth and resource acquisition as effectiveness criteria; and the *internal process model*, in which stability and control are the main issues. Quinn and Rohrbaugh (1983) depict these four models as determined by two dimensions: one that has flexibility and control as its extremes and one that represents an internal versus an external orientation (see Figure 1.2). From this framework additional process indicators of school functioning may be generated.

First of all, it should be noted that the rational goal model does not specify which educational objectives are relevant. Besides knowledge and skills in basic school subjects, other educational aims can be recognised. Two other important categories of educational objectives are *social, emotional and moral development* on the one hand and the development of *general cognitive skills* on the other. For our purposes, these additional categories of educational aims (besides the basic cognitive skills that have been the focus of empirical school effectiveness research) are relevant to the degree that they may require somewhat different teaching approaches and different school organisation arrangements than the process variables that have been shown to matter in the traditional school effectiveness models (Scheerens, 1994).

According to Goodlad (1987), *multi-age* and *inter-age grouping* have the advantage of fostering social and emotional development, apart from being effective in realising traditional education goals. The disadvantages of a competitive achievement-oriented atmosphere are supposed to be mitigated by these organisational arrangements, while the motivational disadvantages of both promotion and non-promotion, as in the graded system, are prevented. The *absence of grading*, and *team teaching*, are seen as measures to realise differentiated adaptive teaching and an integrated, continuous learning route.

Figure 1.2. Typology of effectiveness models
Typologie des modèles d'efficacité



Source: Quinn and Rohrbaugh (1983).

Such approaches are thought to contribute to the degree that students are comfortable and happy in the school.

Educational psychologists increasingly emphasise the importance of self-regulated learning and meta-cognition. "Subject-free" cognitive skills can be acquired in programmes in which learning how knowledge is acquired ("learning to learn") is taught.

The Human Relations Model

The *human relations model* is strongly concerned with the work satisfaction of teachers. Louis and Smith (1990) identified seven "quality of work life indicators":

- *respect from relevant adults*, such as the administrators in the school and district, parents, and the community at large;
- *participation in decision-making* that augments the teachers' sense of influence or control over their work setting;
- *frequent and stimulating professional interaction* among peers (e.g. collaborative work/collegial relationships) within the school;
- structures and procedures that contribute to a high sense of efficacy (e.g. mechanisms permitting teachers to obtain *frequent and accurate feedback about their performance* and the specific effects of their performance on student learning);

- *opportunity to make full use of existing skills and knowledge*, and to acquire new skills and knowledge (self-development); the opportunity to experiment;
- *adequate resources to carry out the job*; a pleasant, orderly physical working environment; and
- *a sense of congruence between personal goals and the school's goals* (low alienation).

Other factors that may contribute to teachers' satisfaction are task differentiation, possibilities of promotion (though these are usually limited), and financial incentives, though this approach, according to some authors, might prove counter-productive (McLaughlin and Yu, 1988). Another point raised by some observers is that it is not clear what these are models of and what they are models for.

The Open System Model

The *open system model* emphasises the responsiveness of schools with respect to environmental requirements. This means, on the one hand, that school organisations can create effective buffers against external threats and, on the other, that schools can manipulate their environments to the degree that their own functioning is not only safeguarded but also improved. In some countries (the Netherlands, for instance), external regulations governing schools have been relaxed and school autonomy enhanced. This state of affairs offers new possibilities but also confronts the school with new requirements, for instance to conduct their own financial policy.

Demographic developments (fewer pupils) may force schools to be active in stimulating student enrolment and "school marketing". Developments in educational technology, initiatives for educational innovations from higher administrative levels, and accountability requirements, can be seen as additional external forces that challenge the school's readiness to change.

In a Dutch study, Gooren (1989) found evidence for a dichotomy between schools that could either cope or not cope with these new external requirements. The schools that could cope more frequently had strong leadership or a collegial structure, in contrast to non- coping schools, which fulfilled the image of the "loosely coupled", segmented school organisation.

The capacity of schools to deal with an increasingly demanding and dynamic environment is described in terms such as "the policy-making potential of school" and "the self-renewing capacity of schools". School organisation characteristics that are thought to contribute to these capacities are:

- leadership (also in the sense of entrepreneurship),
- collegiality,
- capacity for self-evaluation and learning (see, for instance, Morgan's image of the learning organisation in Morgan, 1986, chapter 4),
- overt school-marketing activities,
- strong parental involvement,
- boundary-spanning positions, and
- support of external change agents.

Proxy indicators of the degree of responsiveness are enrolment figures and characteristics of buildings and equipment.

The Internal Process Model

Whereas the human relations model is concerned with social and cultural aspects of "what keeps organisations together", the *internal process model* reflects a preoccupation with formalisation and structure. From this perspective, the following factors are of interest:

- explicit planning documents (such as school curricula and school development plans),
- clear rules regarding discipline,
- formalisation of positions,
- continuity in leadership and staffing, and
- integrated curricula (co-ordination over grades).

Proxy indicators for the stability of school organisations are attendance rates, the number of teaching periods not given, and figures for continuity in staffing.

The ideas for additional process indicators that come from this more comprehensive treatment of organisational effectiveness are summarised in Figure 1.3. (Process indicators induced from the narrower model of school effectiveness research are also included.)

Figure 1.3. Additional topics for process indicators
Autres thèmes à retenir pour l'élaboration des indicateurs de processus

Human relations model

Quality of work life indicators

- respect
- participation in decision-making
- professional interaction
- performance feedback
- opportunity to use skills
- resources
- congruence personal/organizational goals

Internal process model

- planning documents
- disciplinary rules
- management information systems
- formalization of positions
- continuity in staffing and leadership
- integrated curricula
- attendance rates
- lessons "not given"

Open system model

- entrepreneurship
- collegiality
- capacity for self-evaluation and learning
- overt school marketing activities
- parental involvement
- boundary-spanning positions
- external change agents
- student enrolment figures
- resources (buildings, equipment)

Relational goal model

School effectiveness research

- educational leadership
- success-oriented ethos
- monitoring of students' progress
- time on task
- content covered (opportunity to learn)

Broader set of educational goals

- non-gradedness
- team teaching
- individualization, differentiation
- continuous learning route
- time spent on social, emotional, creative and moral development
- "learning to learn" activities
- diagnostic testing

Source: Quinn and Rohrbaugh (1983).

From an examination of Figure 1.3 it is evident that certain features of school functioning are relevant from the perspective of several models, namely leadership, co-ordination among staff, continuity and integration of curricula, and evaluation procedures. Each of the four models puts different emphasis on these common features, however. A second conclusion is that the current set of indicators covers these common features of school functioning reasonably well. The exception is evaluation. Since evaluation reappears in each of the four quadrants of the analysis of multiple organisational effectiveness criteria given in Figure 1.3 (performance feedback, capacity for self-evaluation and organisational learning, monitoring of students' progress and management information systems), there appears to be a compelling argument for the specification of a "summary" indicator relating to the evaluative potential of schools.

Network C decided on the policy of restricting the indicator set developed primarily to the rational goal model, but to add indicators on evaluation practices of schools (all models), user (parents and pupils) involvement in schools (open systems model), teacher qualifications, rewards, working conditions and social status (human relations model), and team stability (internal process model).

The challenges implied in operationally defining, measuring and interpreting these different kinds of process indicators will be discussed in the following sections.

4. Data Collection Strategies

Some of the Network C indicators, namely those that are defined at the system level, have already been calculated. In this chapter, the focus is on the measurement of process indicators defined at school level. Collecting data for process indicators of school functioning poses many problems.

The main reason for this is the fact that the basic data for these process indicators in most cases (*i.e.* countries) are not available in current educational statistics. Nevertheless, descriptive educational statistics remain a relevant source of data for this particular type of indicator. Perhaps, in course of time, it might prove possible to persuade the relevant national statistical offices to start collecting basic data that are relevant for the type of indicator that is being considered here. Examples of such data are: figures on absenteeism, drop-out from schools, figures on the number of lessons "not given" for all kinds of reasons and figures on the stability vs. mobility of teaching personnel.

Apart from educational statistics, three types of data-gathering strategies have been considered:

- the use of *available data sets* from large-scale international comparative surveys;
- "buying into" *international surveys* that are currently being conducted or will take place in the near future; and
- carrying out a *specific INES international survey*.

As to the first alternative, the data sets available from projects of the IEA (International Association for the Evaluation of Educational Achievement) and the IAP (International Assessment of Educational Progress) were thought to be worth analysis. By combining the responses on several relevant items per indicator, it was possible to

examine whether these databases provide acceptable operationalisations of indicator definitions such as "content covered", "time on task", "achievement-oriented ethos" etc. An example from the IEA Second Mathematics Study will illustrate this.

The following variables were used to calculate a composite indicator on *time on task*:

- YAHWKT total homework hours in a typical week (pupil questionnaire)
- TORDERT minutes per week establishing order and discipline
(teacher questionnaire)
- SDAYSYR official school days per year
- SSCHPER average number of periods per school day
- SPERLEN average length of each period in minutes

For calculation of the variable TIONTA (hours per year time on task) the following calculation formula was used:

$$\text{TIONTA} = \left(\frac{\text{SPERLEN}}{60} \times \text{SSCHPER} \times \text{SDAYSYR} \right) - \left(\frac{\text{TORDERT}}{60} \times \frac{\text{SDAYSYR}}{5} \right) + \left(\text{YAHWKT} \times \frac{\text{SDAYSYR}}{5} \right)$$

An interesting possibility of these databases is that process indicators could be directly related to achievement data (while statistically controlling for relevant pupil background characteristics). Major limitations of these databases are that they are restricted to particular subjects (e.g. mathematics and science) and one particular school category (the first phase of secondary education), and that some databases are fairly old (about ten years).

As to the second alternative, "jumping aboard current international surveys" or "piggy-backing" as the English phrase is, four possibilities were considered. The first was the IEA Reading Literacy Study. The database of this study became available in the course of 1993. The second possibility was the Third International Mathematics and Science Study (TIMSS), which is currently being carried out. The third possibility was the IEA project on computers in education. The fourth possibility concerned an international comparative school effectiveness research project (in primary education) abbreviated as ISERP. The countries/territories participating are Australia, Canada, Hong Kong, the Netherlands, Norway, Taiwan, the United Kingdom and the United States.

The conduct of international surveys specifically designed for the INES project is being considered as a real possibility and is currently being tried out in a pilot project. Such a survey should be thought of as a rather brief questionnaire administered to a representative panel of schools in each participating country.

Each of these data strategies has its particular strong and weak points. In addition to the limitations of secondary analyses that have already been mentioned, we should be aware of the fact that in this way one would only be able to produce proxy operationalisations of the indicators concerned. A practical problem with these analyses is that organising the data according to the proper format is usually cumbersome and time-consuming.

"Piggy-backing" on current projects requires much organisation and persuasion of the relevant authorities. Besides, it might prove to be difficult to keep in line with the time sequence these projects have set for themselves. It is still uncertain what the

financial costs would be for adding new items to these international surveys (although these need not necessarily be high).

Designing a specific INES international survey also requires much work: development of instruments, translation of instruments, checking of translations by translating items back into the original language, pilot testing of instruments, sampling, standardised data collection, data storage and retrieval, telecommunications, etc.

An essential precondition for a specific INES survey would be an organisational network of research institutes and/or statistical agencies. In turn, this would require special funding by each of the participating countries.

The experiences that were obtained in 1993 and 1994 with the two first strategies have not been altogether positive. Although several process indicators were computed on the basis of the IEA Reading Literacy Study, the draft versions of these indicators evoked considerable dispute. Several countries, after scrutinising the data generated for them, decided that these were insufficiently valid to be used (e.g. because of ambiguities in the translation of questionnaire items).

The possibilities of "buying into" on-going international surveys also proved to be rather limited. Although at first, through various channels, much energy was put into making the school and teacher questionnaires for the IEA TIMSS project as relevant as possible for the data needs inherent in the indicators conceptualised by Network C, it is not yet clear what the results have been. Apparently, the instrument development went through various cycles, during which internal debate may have forced the data needs of one network of the OECD INES project into the background.

Gradually, the network has come more and more to favour the third strategy: setting up its own data collection. After two instances in which the network collected data at the level of education systems (the locus of decision-making and the teachers and curriculum surveys), in February 1994 the Policy Review and Advisory Group of the INES project approved a survey of schools, designed as a pilot study, to measure process indicators on school functioning.

5. Interpretation and Use of Process Indicators for Policy-Making

As stated elsewhere (Scheerens, 1990), process indicators form a complicated category as far as interpretation is concerned. Following the logic of school effectiveness models, process indicators are to be seen as the antecedent conditions for educational achievement or educational attainment. In this sense, these conditions serve as *auxiliaries* for the interpretation and/or explanation of achievement differences between schools (and perhaps also between countries). Part of the problem is that it is not generally known whether they are sufficient conditions, or even necessary ones.

As mentioned in an earlier section, there is still considerable uncertainty about the knowledge base of school learning. In order to obtain a clear picture of the relevance of process indicators, information is needed of the type: "Increasing the school day by half an hour leads to an expected gain in arithmetic progress of 3/10 of a month" (Levin, 1988). It would be extremely helpful if overviews of factors in education that "work"

(cf. Wang *et al.*, 1993) gave a clear indication of the actual meaning of the effect of process characteristics in terms of increments in students' achievement.

A third possibility of interpreting process indicators would be to take the position that certain process characteristics are valuable "in themselves". For instance, it is not unlikely that panels of educational stakeholders could agree on a statement such as: "A proportion of more than ten percent of lessons not given in a particular subject in secondary education is unacceptable." Despite the danger of *goal displacement*, i.e. the means of education becoming the goals, such an approach is probably the only way to solve the issue of the validity of evaluative process indicators. This would imply that international panels of professionals in education should define attainment levels on each process indicator. If this were to prove impossible for some proposed indicators, these should probably be dropped.

Although this approach – taking certain process characteristics of education as valuable in themselves – is theoretically unrelated to the school effectiveness perspective, in actual practice the two could be combined by assuming that the implicit reasoning in the selection of process indicators would still be based upon the school effectiveness philosophy of "what works". The actual *standard setting* for the indicators, however, would not directly follow from the school effectiveness research but require specific procedures for reaching consensus. Contextual differences of schools in the different countries, to say nothing about processes of human development, should always be taken into account in valuing the process indicators.

Another compelling issue with respect to the application of process indicators is their possible *use in policy-making*. It is striking that a "supply side" approach in selecting indicators appears to have been chosen in the INES project.

The demands of the users of indicators (such as government officers and regional educational authorities), have not been systematically examined as a guiding principle in the selection of the indicators. As pointed out elsewhere (Scheerens, 1990), the relevance of process indicators will depend upon the discretion of decision-makers at various levels of education systems. For instance, it is worth examining what use central offices would have for school-level indicators, in cases where all decision-making authority with respect to school programmes and processes resides with regional offices or even with individual schools. This line of reasoning leads to the conclusion 1) that different information requirements (i.e. different types of indicator) may be required at different administrative levels, and 2) that the relevance to policy of the INES process indicators will not be uniform across countries.

As an illustration of the interpretative problems that may occur when using process indicators derived from available international data sets, some experiences with data collected by the IEA could be mentioned.

From the IEA Reading Literacy Study, some questions referring to the time used for instruction in language and reading and to educational leadership were considered as a basis for indicators. It appeared that in the case of "time", doubts could not be ruled out over interpretation of the relevant questionnaire item. In some countries, language and reading time was computed in terms of the official hours on the timetable, whereas in other countries reading and language during lessons devoted to other subjects (e.g. social

studies and history) were also taken into account. From the data that were available on educational leadership, there was an attempt to make a division between organisational and educational tasks. This attempt was hampered by the fact that some activity areas, most notably "pastoral care", had an educational meaning in some countries and an organisational meaning in others.

It seems that such items from IEA background questionnaires on teachers and schools sometimes do not survive the heavy scrutiny that is employed when the attempt is made to use them as a basis for national indicators.

6. Conclusion

Indicators of educational programmes and processes are among the "avant-garde" categories of indicators that are being explored in the INES project. The international network that is dealing with this category of indicators has reached a consensus on accepting a broad perspective on the organisational effectiveness of schools as the general rationale for selection. As a matter of course, this does not solve all problems of measurement, interpretation and use of this category of indicators for policy-making. With respect to these issues, pervasive problems remain for the network to tackle. These are:

- the question of evaluative interpretation and standard setting;
- the question of how to reconcile the rather heavy data-collection requirements of the school-level indicators, with the general idea of relatively easy and global measurement, which is one of the defining characteristics of indicators;
- the question of providing appropriate contextual information with the actual calculations, in which national differences will have to be taken into account;
- the question of ensuring the relevance to policy-making of the process indicators; here too, differences between countries, particularly with respect to the distribution of formal authority between levels of the education system, will have to be taken into account.

The third phase of the INES project has been a period in which quite a few experiences were obtained in tackling these major problems. Although definitive conclusions have not been drawn yet, the opportunities for setting up data-collection strategies that are efficient and adapted to the real data needs of the OECD in this area seem to be more promising and should be further investigated. They would supplement the use of available international data sets such as those of the IEA or IAEP studies.

References

- GOODLAD, J.I. (ed.) (1987), "The ecology of school renewal", University of Chicago Press, Chicago.
- GOOREN, W.A.J. (1989), "Kwetsbare en weerbare scholen en het welbevinden van de leraar", in J. Scheerens and J.C. Verhoeven (eds.), *Schoolorganisatie, beleid en onderwijskwaliteit*, Lisse, Swets and Zeitlinger.
- LEVIN, H.M. (1988), "Cost-effectiveness and educational policy", *Educational Evaluation and Policy Analysis*, Vol. 10, pp. 51-70.
- LOUIS, K.S. and SMITH, B.A. (1990), "Teachers' work: current issues and prospects for reform", in P. Reyes (ed.), *Productivity and Performance in Educational Organisations*, Newbury Park, Sage, pp. 23-47.
- MCLAUGHLIN, M.W. and YU, S. (1988), "School as a place to have a career", in A. Liebermann (ed.), *Building a Professional Culture in Schools*, New York.
- MORGAN, G. (1986), *Images of Organisations*, Beverly Hills, Sage.
- OECD (1991), *Draft Technical Handbook for Education Indicators*, Paris.
- OECD (1992), *Education at a Glance: OECD Indicators* (bilingual), CERI, Paris.
- QUINN, R.E. and ROHRBAUGH, (1983), "Spatial model of effectiveness criteria towards a competing values approach to organisational analysis", *Management Science*, Vol. 29, pp. 363-377.
- SCHEERENS, J. (1990), "School effectiveness research and the development of process indicators of school functioning", *School Effectiveness and School Improvement*, Vol. 1, pp. 61-80.
- SCHEERENS, J. (1992), *Effective Schooling. Research, Theory and Practice*, Cassell, London.
- SCHEERENS, J. (1994), "The school-level context of instructional effectiveness: a comparison between school effectiveness and restructuring models", *Tijdschrift voor Onderwijsresearch*, Vol. 19 (1), pp. 26-38.
- WANG, M.C., HAERTEL, G.D. and WALBERG, H.J. (1993), "Toward a knowledge base for school learning", *Review of Educational Research*, Vol. 63, pp. 249-294.

Chapitre/Chapter 2

**Comparer les structures de décision des systèmes éducatifs :
un bilan de l'approche quantitative**

***Comparison of Decision-Making Structures in Education Systems:
A Review of the Quantitative Approach***

par

Denis Meuret

IUFM de Bretagne, Rennes, France

et

Jacques Prod'hom et Eugen Stocker

Office fédéral de la statistique, Berne, Suisse

Les auteurs ont réalisé, dans le cadre du projet OCDE/CERI sur les indicateurs internationaux des systèmes d'enseignement (Projet INES), une étude quantitative des prises de décision dans les systèmes éducatifs de 14 pays (Étude quantitative des décisions, EQD). Cette étude a permis de calculer des indicateurs pour *Regards sur l'éducation* (OCDE, 1993) et de rédiger un rapport publié par le CERI (OCDE, 1995). Elle se fonde sur un questionnaire rempli par des représentants des pays et portant sur 35 décisions. Son objet est la répartition des décisions entre l'établissement et les trois niveaux administratifs qu'on peut trouver au-dessus de lui (local, régional et national), chaque décision étant en outre repérée par le domaine auquel elle appartient et par la façon dont elle est prise (en autonomie, en concertation, dans un cadre fixé). Ce texte considère cette étude dans une double perspective, d'une part celle de tentatives précédentes pour comparer l'organisation des systèmes éducatifs, d'autre part en essayant de mettre en parallèle ses résultats avec certaines caractéristiques du coût et de l'efficacité de l'éducation dans les pays correspondants.

Les études précédentes évoquent d'assez nombreux domaines de décision en matière d'éducation, absents de l'EQD, et qui pourraient être intégrés dans une nouvelle version du questionnaire. La structure de décision du système a été mise en rapport avec d'autres caractéristiques des systèmes éducatifs : la réussite à la fin de l'enseignement secondaire (R5), le niveau de lecture à 14 ans (R1), et un indicateur de coût, les dépenses par

élèves (P6). Les comparaisons rudimentaires effectuées ne mettent en évidence aucune relation entre la structure de décision et les autres caractéristiques. Tout au plus, laissent-elles ouverte la possibilité d'une relation négative (légère) entre l'importance des décisions prises au niveau national et le fait d'obtenir un diplôme à la fin de ses études secondaires, et d'une relation positive (légère) entre les prérogatives de l'ensemble «établissement + niveau local» et le coût du système.

*

* *

Summary

As part of the OECD CERI project on education indicators (INES), the authors made a quantitative study of decision-making (QSD) in the education systems of 14 countries. From this study, it was possible to calculate indicators for Education at a Glance (OECD, 1993) and to produce a report published by CERI (OECD, 1995). The study was based on a questionnaire on 35 types of decision, which was completed by respondents in the countries. The aim was to reveal the division of decision-making between establishments and the three administrative levels (local, regional and national) which are above them. Each decision was also categorised by the domain to which it relates and the way in which it is taken (autonomously, in consultation, or within a fixed framework).

This chapter considers the study from two angles: by placing it in the context of previous attempts to compare the organisation of education systems; and by contrasting it with certain characteristics of the cost and effectiveness of education in the countries involved, which are described by other indicators in Education at a Glance.

Earlier studies were made in the United States in the 1930s and again in 1960; by the International Bureau of Education of UNESCO in the late 1970s; and then by the European Community in the late 1980s. Their political and intellectual background, their methods and their results are compared with those of QSD.

Until the mid-1970s, the studies were conducted against a background of widespread political and intellectual acceptance of centralisation as inevitable and representative of progress. They also used questionnaires, although less detailed than QSD, and their approach was legalistic rather than organisational. They ignored the educational establishment as a potential level of responsibility, while QSD stresses the school's point of view – perhaps to excess.

Most of the earlier studies, like QSD, conclude that the opposition between centralised and decentralised systems is oversimplified, but they often replace this distinction by a division into only three groups: systems managed at the local, regional or national level. As QSD makes plain, this division is rendered inadequate by the growing role of

the school, the scale of the delegation of responsibility from one level to another, and the large number of decisions taken by several levels at once.

The earlier studies call attention to a number of domains of educational decision-making which are absent from QSD and should be included in any new version of the questionnaire. Among these are the levels at which decisions are taken on teacher training and recruitment, and on other legal and financial aspects of education.

The decision-making structure is also compared with other characteristics of the education system: upper secondary graduation (R5 in Education at a Glance, OECD, 1993), performance in reading at the age of 14 (R1), and an indicator of cost, expenditure per student (P6).

Crude comparisons show no evidence of a link between the decision-making structure and the other characteristics, particularly students' level of reading. At most, they leave open the possibility of a small negative relationship between the proportion of decisions taken at the national level and upper secondary graduation, and a small positive relationship between responsibility at "school + local level" and the costs of the system.

The interest of these comparisons lies rather in their methodology. They suggest how a cross-study of two indicators may be conducted, provided attention is given to the comparability of populations, and parts of the system.



1. Introduction

Des données sur le processus de décision dans le système éducatif ont été recueillies auprès d'une quinzaine de pays, dans le cadre du projet OCDE/CERI sur les indicateurs de l'enseignement, au sein du Réseau C sur les caractéristiques du fonctionnement des systèmes éducatifs. Elles ont permis de calculer cinq indicateurs expérimentaux dans l'édition de 1992 de *Regards sur l'éducation* (P17 à P21) et quatre indicateurs (P18 à P21) dans l'édition de 1993. Elles ont été analysées de façon détaillée par les auteurs dans un rapport réalisé pour le CERI (OCDE, 1995).

L'occasion est ici donnée de mettre davantage ce travail en perspective en le comparant à d'autres descriptions du partage des responsabilités dans les systèmes éducatifs. On tentera de mettre ainsi en évidence quelques-uns de ses avantages et quelques-unes de ses limites. Dans ce texte, l'ensemble du travail – l'enquête, le calcul des indicateurs et la réalisation du rapport – sera désigné par Étude quantitative des décisions (EQD).

On compare ensuite certains indicateurs issus de ces données à d'autres indicateurs de *Regards sur l'éducation* traitant du coût ou de l'efficacité des systèmes éducatifs. La quantification permet en effet ce type de mise en relation. Malgré le faible nombre de pays pour lesquels l'exercice est possible, elle aura surtout une vertu méthodologique :

commencer à repérer que la mise en relation du coût, des processus et des résultats requiert des indicateurs qui traitent de chacun des trois domaines.

Cette double mise en perspective devrait être utile à l'amélioration de la construction des données dans un domaine important pour la conduite des systèmes éducatifs.

2. Autres approches de l'organisation des systèmes éducatifs

La comparaison de l'organisation des systèmes éducatifs est une préoccupation déjà ancienne. En 1962, le Department of Education and Welfare américain publiait un ouvrage intitulé *Ministries of Education, their functions and organisations* (Heath, 1962), illustrant ainsi que les relations internationales étaient une prérogative du niveau fédéral. Ce travail avait mis à contribution les ministères de nombreux pays ainsi que leurs ambassades à Washington. Il décrivait :

- les organismes qui ont la responsabilité générale de l'éducation, le mode de nomination de leurs dirigeants;
- la distribution de l'autorité, des responsabilités, du financement au sein des systèmes éducatifs; ainsi que
- le mode d'administration, à travers la description des voies qu'empruntent dans chaque pays, tant les suggestions aux plus hauts responsables du système que les instructions ou recommandations que ceux-ci envoient au « terrain ».

Cet ouvrage mettait à jour une publication précédente, réalisée en 1930. Il est tentant d'observer que, dans le cadre des indicateurs internationaux de l'enseignement, le CERI a également repris cette question trente ans plus tard. Il faut cependant aussi tenir compte des travaux que l'UNESCO, à travers le Bureau international de l'éducation (BIE), puis la Communauté européenne, grâce à l'unité européenne d'Eurydice, ont réalisés dans ce domaine, respectivement vers la fin des années 70 et vers la fin des années 80.

À l'occasion de la Conférence internationale de l'éducation de 1977, le BIE a demandé aux pays membres de rédiger deux rapports, l'un décrivant l'état actuel du système éducatif, l'autre son évolution depuis deux ans. Le projet était de relancer l'Annuaire international de l'éducation, lequel avait cessé de paraître en 1969, par la publication alternée d'un rapport « dynamique » et d'un rapport « statique », chacun paraissant tous les deux ans. Analysés par d'éminents spécialistes de l'éducation comparée, les rapports nationaux ont donné lieu à deux ouvrages (Holmes, 1979; Cowen, 1982). Leur parenté est certaine avec la tentative de cerner les processus de décision dans le cadre des indicateurs de l'enseignement, même si les voies utilisées sont différentes.

En principe, le champ couvert par cet annuaire est semblable à celui des indicateurs internationaux de l'enseignement dans leur ensemble. Toutefois, une série de raisons fait qu'il se réduit à l'administration des systèmes éducatifs, de sorte qu'il correspond assez bien aux indicateurs des processus de décision.

Le réseau Eurydice, quant à lui, a constitué des « dossiers nationaux » sur les structures des systèmes éducatifs de chacun des pays membres de la Communauté européenne, alors au nombre de douze. Une partie de ces dossiers a permis de rédiger des fiches synthétiques sur la répartition des compétences dans le domaine de l'éducation.

Cette synthèse a été publiée (Eurydice, 1991). Comme l'EQD, elle laisse de côté l'enseignement pré-primaire et l'enseignement supérieur.

A notre connaissance, il s'agit là des dernières approches générales du fonctionnement des systèmes éducatifs avant l'EQD. Cependant, dans le cadre de rapports thématiques – qu'il s'agisse par exemple des programmes scolaires (Skilbeck, 1990) ou des examens (Eckstein et Noah, 1993) – la question de la répartition des compétences en la matière a depuis été traitée plusieurs fois.

Dans ce chapitre, l'environnement, la méthode ainsi que les résultats de ces travaux seront confrontés, trop globalement sans doute, à ceux de l'étude quantitative des décisions (EQD), avant que l'on n'en tire des leçons pour la suite de ce travail.

Un environnement politique et intellectuel différent

L'environnement politique

En octobre 1967, Alexander King, alors Directeur des Affaires scientifiques de l'OCDE, prononce un discours à la Conférence internationale sur la crise de l'éducation (King, 1967). Cet exposé transporte le lecteur d'aujourd'hui dans un monde dont il a oublié même qu'il ait pu exister.

Ce monde est celui du progrès, conçu comme une tendance historique irrépessible, venue de partout et de nulle part en particulier, et dont la rapidité oblige tout un chacun à s'adapter sous peine d'en manquer les innombrables possibilités.

Comme chacun sait, ce monde pressé et optimiste a été remplacé par un autre, aussi pressé mais moins optimiste, dans lequel le progrès procède de la concurrence, alors qu'auparavant, c'était plutôt la concurrence qui procédait du progrès, le gagnant étant celui qui se dotait des structures le plus à même de dominer le changement et de maîtriser ce progrès.

Pour donner une image, on pourrait dire que dans le monde « d'avant cette crise de l'éducation », la compétition était pensée plus volontiers sur le modèle de la course à pied que sur celui du match de boxe. Les plus forts impulsaient un mouvement qui profitait à tous par le jeu de l'émulation. Aujourd'hui, le gagnant serait plutôt celui qui a éliminé ses concurrents et reste seul maître du champ de bataille. Pour une part, il s'agit là de différences rhétoriques, mais dont chacun peut ressentir les conséquences.

Dans le monde « d'avant la crise », il était évident que les systèmes éducatifs seraient de plus en plus gérés au centre, par des professionnels compétents :

« Dans de nombreux pays, l'enseignement est encore fortement décentralisé, bien qu'il existe une tendance générale vers une direction centrale par l'État. [Cette tendance] est probablement inévitable parce qu'on se rend compte de plus en plus des rapports existants entre l'enseignement et le développement économique et social (King, 1967, p. 1). »

La décentralisation, c'est la matrice – « Jusqu'à une date récente, les administrations de l'Instruction publique à travers le monde ont lentement pris forme, souvent à partir d'éléments locaux (...) » – qu'il importe de quitter :

« Dans les pays où l'éducation est décentralisée, on est mal renseigné sur la manière dont les décisions sont prises, sur ceux qui décident, les pressions exercées et les convictions qui les motivent. Il n'est pas possible d'être sûr que ceux qui dirigent les écoles soient suffisamment au courant des objectifs nationaux, dans leur ensemble et dans leur détails, pour être à même de prendre, au plan local, des décisions en accord avec la politique nationale (King, 1967, p. 6). »

Ce mouvement vers une plus grande implication de l'État s'inscrit d'ailleurs dans une tendance longue dont Heath (1962, p. 137) se fait l'écho : « Indicative of the trend toward greater government responsibility for education are specific provisions on education in the majority of the operative Constitutions which entered into force during the 20th century and, particularly, around the mid-20th Century. »

L'évidence aujourd'hui est plutôt inverse et, dans de nombreux pays, au cours des années 80, les niveaux locaux ou régionaux ont accru leurs prérogatives au détriment de l'État central. Deux remarques cependant doivent amener à tempérer l'opposition entre ces deux mouvements.

D'abord, le modèle de centralisation prôné par King en 1967 n'est pas celui que critiquent aujourd'hui les tenants du modèle local, mais celui de la gestion scientifique des grandes entreprises. Il écrit par exemple : « L'éducation est ainsi la dernière des grandes industries à n'avoir pas été atteinte par la révolution scientifique » (King, 1967, p. 2). On retrouve cette référence dans Haag (1981), cette fois au service de la décentralisation, le modèle étant la gestion décentralisée des grandes firmes. Rétrospectivement, on s'aperçoit que les experts ont longtemps proposé de gérer les systèmes éducatifs avec les techniques de la planification décentralisée. Ce modèle exigeait à la fois plus et moins de centralisation que les organisations antérieures. Mais il s'est trouvé invalidé par l'émergence du marché comme « discours de gouvernement » à la fin des années 70. Au ballet des symboles, les *bureaucrates routiniers* ont alors remplacé les *experts soucieux du progrès* comme figures du modèle centralisé, et les *petits entrepreneurs innovants* ont succédé aux *notables, conservateurs et sensibles aux pressions contraires à l'intérêt général*, comme figures du modèle local.

Ensuite, il n'est pas sûr que le mouvement annoncé par King ne se soit pas produit, en quelque sorte silencieusement, recouvert par la rhétorique du modèle local. Dans plusieurs pays, l'accroissement du pouvoir de décision des établissements scolaires peut être compris comme un renforcement du pouvoir des professionnels au détriment des administrateurs locaux ou centraux. Ailleurs, à travers la définition des programmes (Skilbeck, 1990), des examens (Eckstein et Noah, 1993), ou encore à travers la mise en œuvre de dispositifs d'évaluation, le pouvoir central a accru son influence, sans pour autant prendre davantage de décisions qu'auparavant. Ceci pourrait d'ailleurs inciter à mesurer l'influence des différents niveaux de décision sur les systèmes éducatifs en utilisant également d'autres critères que la proportion de décisions prises.

Ces différences dans les « discours de gouvernement » ne sont pas immédiatement perceptibles dans les études des différentes périodes, que ce soit en faveur de la centralisation ou de la décentralisation, pour autant que les auteurs de ce texte puissent en être de bons juges. Elles agissent pourtant à travers ce qu'on appellera, faute d'avoir trouvé mieux, l'environnement intellectuel.

L'environnement intellectuel

L'étude quantitative des décisions (EQD) et les études précédentes portent les marques de deux caractéristiques de leurs environnements intellectuels respectifs : la primauté, aujourd'hui, de l'espace sur le temps dans la pensée de l'organisation, et celle de la décision sur la planification.

a) L'espace, le temps

Dans le monde d'après la crise, la compétition est davantage visible, puisqu'elle est valorisée, en tant que source de progrès. Il en résulte, parmi d'autres caractéristiques plus inquiétantes, que la comparaison spatiale (entre pays, systèmes éducatifs, établissements scolaires, etc.) prévaut sur la comparaison temporelle, longitudinale. Non seulement elle est pratiquée plus souvent, mais encore, quand les deux approches sont combinées (si l'on compare l'évolution de deux pays, par exemple), elle domine : on compare deux pays en prenant l'évolution générale pour critère, plutôt que d'appréhender l'évolution générale en prenant les pays pour critères.

Les comparaisons internationales s'efforçaient de trouver des invariants, ou la tendance générale qui pouvait représenter le mouvement du progrès. L'analyse consistait plutôt à situer chaque pays vis-à-vis de cette tendance centrale. L'idée qui sous-tend l'EQD est plutôt de quantifier l'approche du processus de décision de façon à chercher si, à terme, certaines structures permettent un meilleur enseignement que d'autres.

La situation optimale n'est plus celle qui témoigne d'une bonne adaptation à l'histoire mais plutôt d'une bonne adaptation à la configuration propre de chaque pays et à la compétition avec les autres.

b) La décision, la planification

Les études qui ont précédé l'EQD traitent en général, non des décisions, mais de la responsabilité. On y trouve des énoncés tels que celui-ci : dans les systèmes centralisés, « le ministère de l'Éducation est juridiquement responsable de l'orientation, de la supervision et de l'inspection effective de l'ensemble des tâches administratives relatives à l'Éducation » (Cowen, 1982, p. 63). Il peut d'ailleurs arriver que l'objet de la responsabilité soit assez peu précisé : dans tel pays, « l'administration de tout le système éducatif relève du ministère de... Ce sont toutefois les conseils scolaires provinciaux qui sont responsables au niveau des provinces et les conseils scolaires municipaux au niveau des municipalités » (Holmes, 1979, p. 159). Les auteurs des études regrettent parfois que le cadre de leur travail (l'exploitation de rapports nationaux réalisés par des administrateurs) les cantonne dans cette approche juridique. Ils se prennent à rêver d'une étude détaillée dont on peut imaginer qu'elle ressemble à celle qui a été mise en œuvre pour les indicateurs internationaux de l'enseignement : « Les systèmes éducatifs nationaux sont fréquemment administrés par le truchement d'organismes officiels aux niveaux national, provincial et local. Ces divers échelons se partagent les tâches, et seule une recherche comparative minutieuse permettrait de montrer comment ces rapports diffèrent d'un pays à l'autre » (Holmes, 1979, p. 13). En vérité, il est probable que ces auteurs envisagent une

étude tellement détaillée qu'elle «devrait durer plusieurs années», ce qui la rend irréalisable.

Cependant, il est intéressant de noter qu'ils ne semblent pas trop déçus, car leur idéal, cohérent avec ce que King appelait une gestion scientifique des systèmes éducatifs, repose sur une étude des politiques mises en œuvre pour l'éducation, et non des tâches ou des décisions. Plutôt que d'indiquer qui a la responsabilité de tel domaine, il faudrait indiquer 1) qui construit, 2) qui adopte, et enfin 3) qui met en œuvre chaque politique (Holmes, 1979, p. 14). Bref, il s'agirait de rendre compte de cette planification décentralisée qui a été évoquée plus haut.

Il faut insister sur ce point : ce passage de la planification aux décisions, d'une conception «scientifique» de la gestion publique à une conception orientée vers le décideur n'a rien de mineur, ni d'anecdotique. Il nous ramène à la différence entre gestions orientées vers la maîtrise du temps plutôt que de l'espace. Il illustre le passage d'un univers où la planification est prisee parce qu'on y voit le moyen de dominer la complexité d'une évolution de plus en plus rapide, à un autre où la décision est privilégiée, parce que ce qui compte est davantage la capacité de mobiliser à un moment donné et de façon optimale des ressources sur un certain territoire.

Une autre raison éloigne les auteurs précédents de l'étude des décisions (EQD) telles que la réglementation les distribue et qui sont en quelque sorte «l'unité de travail» de l'EQD. La sociologie des organisations venait de procéder à une relativisation sévère des approches juridiques (cf. par exemple, Crozier, 1963). Elle montrait que le fonctionnement réel des organisations n'était pas celui que la réglementation prévoyait, et donc que l'influence qu'exerce une instance pouvait fluctuer de façon considérable sans que changent les textes juridiques qui définissent ses compétences. Ainsi, selon Ranson (1985), en Angleterre, l'équilibre des pouvoirs entre le Department of Education and Science et les Local Education Authorities (LEA) a-t-il changé plusieurs fois depuis la loi fondatrice de 1944, sans que la nature de leurs compétences respectives ait été juridiquement modifiée, mais, sans doute, parce que l'importance et la pertinence de leurs compétences changeaient en fonction de l'évolution historique.

Des approches différentes

Les différences d'approche entre l'EQD et les travaux qui l'ont précédée tiennent surtout au degré de formalisation de l'investigation, puis au traitement de l'établissement scolaire d'une part et des compétences déléguées d'autre part.

Degré de formalisation

La structuration de l'interrogation est moins exhaustive, moins générale, plus proche de l'activité d'enseignement elle-même, dans l'EQD que dans les travaux précédents.

On se souvient que l'EQD répartit les décisions en quatre domaines :

- organisation pédagogique (décisions sur les parcours au sein du système, les méthodes et manuels pédagogiques, l'établissement d'accueil, l'évaluation, l'organisation de l'aide au travail scolaire);

- planification et structures de l'enseignement (construction des formations, des programmes, des diplômes, délimitation des établissements et des classes);
- gestion du personnel (recrutement, salaires, conditions de travail, carrière);
- affectation et utilisation des ressources (humaines et financières).

Holmes (1979) utilise les catégories suivantes : buts généraux, administration (incluant la répartition des compétences), financement, structure et organisation (en un sens plus large que l'EQD, puisqu'il s'agit surtout des filières et des types d'établissement), programmes d'études, formation des enseignants. Des renseignements analogues à ceux de l'EQD se trouvent dans toutes ces catégories, et pas seulement dans celle de l'administration, parce que, faute de pouvoir y disposer de renseignements statistiques sur tel ou tel domaine, l'auteur indique en fait les instances qui sont compétentes en la matière. A cet égard, il est clair que l'insertion de l'EQD dans l'ensemble du projet des indicateurs de l'enseignement a permis de délimiter plus précisément son objet.

Eurydice utilise les catégories suivantes : législation globale et administration (incluant certaines questions de répartition des compétences), structure du système d'enseignement (incluant des éléments sur la structure générale telle que traitée par Holmes, mais aussi les programmes, l'évaluation et les diplômes), financement (qui finance?).

Toutefois, la différence entre l'EQD et les travaux précédents réside surtout dans le degré de formalisation et de détail avec lequel l'enquête est menée.

Dans Heath (1962) et dans Holmes (1979), comme d'ailleurs dans Eurydice (1991), les limites de l'information disponible sont telles que les auteurs regrettent parfois de devoir se contenter d'une description. Il arrive que le lecteur apprenne surtout le nom des institutions compétentes, ce qui est très utile pour une visite professionnelle, mais relativement peu pour une analyse. Ainsi peut-on trouver dans Heath (1962) la liste des ministères responsables de quelques écoles en France, sans que leur importance vis-à-vis du ministère de l'Éducation nationale soit précisée. Il arrive aussi que la nature des compétences des différents niveaux ne soit pas indiquée de façon très précise. On apprend dans Holmes (1979, p. 138) qu'en République démocratique allemande, « les autorités locales ont une grande liberté d'action dans les affaires de caractère local ».

Si le CERI (OCDE, 1995) évite ce type d'énoncés, c'est uniquement parce que le cadre de l'enquête était beaucoup plus contraignant. L'EQD a reposé sur le pari que, au-delà des différences entre les systèmes éducatifs, certaines fonctions leurs étaient communes : recruter des enseignants, décider ce qu'on enseigne en classe, ouvrir ou fermer un établissement, etc. On pouvait dès lors se demander où et comment ces fonctions étaient remplies. Cette démarche impliquait une approche en réalité plus gestionnaire que juridique, ou plutôt que l'on posât une question juridique (qui est compétent en la matière?) à propos d'objets (les décisions) dont la délimitation avait cherché à épouser une logique de gestion.

Comme dans les approches précédentes, cette démarche a ses points faibles, mais ils apparaissent davantage dans l'EQD précisément à cause de son plus grand degré de précision et, partant, d'ambition. Par exemple, comment s'assurer que la liste des fonctions est suffisamment complète, comment attribuer aux différentes fonctions un poids

proportionnel à leur importance ? Et selon quels critères, selon quels juges, déterminer cette importance ?

L'EQD rencontre d'autre part des obstacles liés à ce plus grand degré de précision. Il arrive que l'organisation d'un système prive de pertinence telle décision. Par exemple, en Norvège, en CITE 1 et 2, les élèves passent automatiquement dans la classe supérieure d'un enseignement de tronc unique. La question de savoir qui décide du passage dans la classe supérieure (P2) est donc dénuée de pertinence pour ce pays, qui n'y a évidemment pas répondu. L'étude gère ce problème techniquement, en accroissant le poids des autres décisions du domaine. Cette solution suppose cependant que ce problème ne se pose pas trop souvent pour un pays donné, ce qui a été le cas dans la plupart des pays qui ont répondu à l'enquête, dénotant une certaine homogénéité des systèmes scolaires. Cependant, il se peut que les pays qui n'ont pas répondu à l'enquête s'en soient abstenus précisément parce qu'ils anticipaient des problèmes de cet ordre.

L'établissement scolaire

La différence la plus claire entre les approches des années 60 (Heath), des années 70 (Holmes, Cowen), et celle de l'EQD est l'absence dans les premières du niveau de l'établissement scolaire.

Cette différence est d'autant plus nette que l'établissement scolaire est en quelque sorte l'objet privilégié de l'EQD, dans la mesure où la question traitée est en réalité la suivante : « Où sont prises les décisions que l'établissement serait susceptible de prendre ? »

Cette différence a plusieurs origines, l'une étant que les prérogatives des établissements scolaires se sont accrues dans la plupart des pays depuis 1980. Ne pas tenir compte de l'établissement scolaire serait, au moins aujourd'hui, une lacune grave pour l'analyse.

Cependant, cette différence attire l'attention sur une limite des comparaisons faites, à partir de l'EQD, sur les prérogatives des niveaux autres que l'établissement. Si par exemple on veut comparer le poids respectif des régions, il faudrait le faire à partir des décisions qu'elles-mêmes seraient susceptibles de prendre et pas seulement à partir des décisions que l'établissement serait susceptible de prendre. Si par exemple un établissement ne peut jamais fixer la durée de l'année scolaire, une région peut par contre en décider (c'est le cas des cantons suisses). Or, cette décision ne figure pas dans le questionnaire de l'EQD, de sorte que les prérogatives des régions qui la prennent sont sous-estimées.

La délégation de responsabilités

Quel niveau est responsable de la décision prise régionalement par un échelon décentralisé de l'administration nationale ? Lorsqu'un niveau se dessaisit lui-même d'une prérogative au profit d'un autre, auquel des deux faut-il l'attribuer ? Pour apprécier l'importance d'un niveau de décision dans un système éducatif, faut-il tenir compte de la nature juridique du texte qui lui confère ses prérogatives ?

Dans les études précédentes, sont tenues pour équivalentes les décisions qui sont prises au niveau de l'administration centrale ou dans des échelons décentralisés de l'État.

Ce n'est pas le cas dans l'EQD. La raison «matérielle» en est que les pays, invités à préciser dans leurs réponses si les administrations locales concernées étaient ou non autonomes, l'ont rarement fait, de sorte qu'il a fallu ignorer cette distinction. La raison de fond en est que le critère de proximité de l'établissement scolaire a paru supérieur à celui de la légitimité politique.

Lorsqu'une compétence est déléguée par un niveau à un autre, l'EQD l'attribue au niveau destinataire. Les études précédentes, parce qu'elles travaillent sur la responsabilité plutôt que sur la décision, l'attribuent en général au niveau qui a délégué. Enfin, ces études mentionnent souvent la nature du texte juridique qui attribue la responsabilité : Constitution, Loi, simple décret, etc.

Tout ceci manifeste encore la prééminence dans l'EQD de la gestion sur le droit. En droit, l'instance pertinente est plutôt celle qui institue. Il faut par exemple écrire que dans tel pays, «on entend par enseignement public l'enseignement institué par le pouvoir central ou par les pouvoirs locaux, c'est-à-dire par l'État ou par les communes» (Cowen, 1982, p. 94). Non seulement une administration doit être considérée comme appartenant au niveau qui lui confère sa légitimité politique, au niveau où se trouve l'assemblée élue dont elle procède, mais encore la légitimité de cette assemblée varie selon la nature du texte qui l'a instituée. En droit, les conseils régionaux (élus) français, qui détiennent leurs attributions éducatives de la loi, sont moins autonomes que les *Länder* allemands, qui les détiennent de la Constitution. Ceux-ci, à leur tour, sont sans doute moins autonomes que les États américains, qui les détiennent du 10^e amendement de la Constitution prévoyant que tous les pouvoirs appartiennent aux États, sauf ceux qui sont explicitement confiés au niveau fédéral par la Constitution, ce qui n'est pas le cas de l'éducation. Pour qui s'intéresse à la possibilité ou aux moyens de modifier la structure de décision dans un pays, de telles considérations sont importantes.

En réalité, l'EQD ignore la suprématie des textes instituant l'autorité avec laquelle telle ou telle instance exerce ses prérogatives tout comme l'étude évite soigneusement de prendre en compte aucun des très nombreux facteurs qui viennent «pondérer» cette autorité. Ceci la distingue d'une approche juridique comme d'une approche sociologique. Il s'agit d'une approche organisationnelle, sans aucun doute moins ambitieuse ou moins «pure» (elle demande aux textes juridiques d'indiquer les compétences de chacun, sans aller jusqu'au terme d'une logique juridique) que les deux autres. Sa légitimité est moins analytique que politique. Rien ne permet de dire que l'approche par l'organisation est supérieure à l'approche par le droit ou par la sociologie, mais c'est l'organisation que les responsables des systèmes éducatifs ont transformée, par une nouvelle distribution juridique des compétences, et c'est cette transformation que l'EQD voudrait aider à évaluer.

Les résultats

Les données fournies par l'EQD sont beaucoup plus détaillées que celles dont disposaient les auteurs précédents. Elles distinguent 34 décisions, alors qu'environ une demi-douzaine de «domaines de responsabilité» apparaissent dans les autres études. Il est donc évident que l'analyse peut être plus précise, que ce soit par la distinction de plusieurs modes et domaines de décision, par le plus grand détail des thèmes traités ou par la quantification. Celle-ci a permis de ne pas se cantonner à une approche typologi-

que, de rechercher par exemple des corrélations entre le poids de différents niveaux dans la structure de décision, de décrire une situation moyenne qui pouvait ensuite servir de référence, etc.

Mais il est également évident qu'il y a peu d'intérêt à souligner ce surcroît de précision. Il est plus intéressant d'étudier la façon dont l'EQD d'une part, les études précédentes d'autre part, prennent leur distance avec l'opposition répandue entre systèmes éducatifs centralisés et décentralisés.

Dans tous les rapports étudiés, comme dans celui du CERI (OCDE, 1995), l'opposition entre systèmes centralisés et décentralisés est tenue pour insuffisante. Mais c'est en général pour proposer, comme Cowen (1982), trois catégories de systèmes éducatifs : centralisés, régionaux, locaux, selon le niveau où est exercé – il vaudrait mieux dire, d'où procède – l'essentiel des « fonctions » ou des « responsabilités ».

En 1977, puisque ses données datent de cette année-là, parmi les pays de l'OCDE, Cowen tient pour « centralisés » l'Espagne, la France, le Japon, la Suède, la Turquie. Il tient pour « régionaux » l'Australie, le Canada, les États-Unis, l'Allemagne et pour « locaux » la Norvège, les Pays-Bas et le Royaume-Uni.

Si l'on classe les pays avec les données de l'EQD, en fonction de l'importance respective des trois niveaux de décision situés au-dessus de l'établissement, comme l'a fait Cowen, on obtient, pour le niveau CITE 2 public, les résultats suivants. Les pays « centralisés » sont : l'Autriche, l'Espagne, l'Irlande, la Nouvelle-Zélande, les deux derniers ayant par ailleurs des établissements très autonomes. Les pays « régionalisés » sont la France et la Suisse, mais il faut ajouter que le niveau national prend en France une seule décision de moins que le niveau régional. Les systèmes « locaux » sont ceux de l'Allemagne, de la Belgique, du Danemark, des États-Unis, de la Finlande, de la Norvège et de la Suède.

Par l'effet combiné des évolutions survenues entre 1977 et 1990 (en particulier en France et en Suède), des différences d'approches et de la plus ou moins grande précision des données, seules l'Espagne et la Norvège se trouvent placées dans la même catégorie par Cowen et par l'EQD.

On peut rappeler ici que la typologie issue de l'EQD qui mobilise le maximum de dimensions (OCDE, 1995) distingue surtout trois groupes de pays : l'Irlande, le Portugal et la Nouvelle-Zélande, qui se caractérisent par une faible intervention des niveaux local et régional ; l'Autriche, l'Espagne et la France, qui se distinguent par un poids à peu près équivalent du niveau national, des deux niveaux intermédiaires pris ensemble, et de l'établissement ; et enfin, les États-Unis, l'Allemagne et les pays nordiques, dans lesquels l'établissement et le niveau local prennent l'essentiel des décisions.

Ces différences sont dues, pour une part, aux diverses approches relevées ci-dessus.

Par exemple, Cowen signale que certains États américains délèguent aux districts une partie de leurs compétences, mais tant ses données que son approche (qui tient les compétences déléguées comme appartenant encore à l'instance qui a choisi de s'en dessaisir) l'empêchent de saisir l'ampleur réelle de ce mouvement de délégation. Il en va probablement de même pour l'Allemagne.

D'autre part, l'EQD montre qu'il y a souvent plusieurs niveaux dominants et pas un seul, ce que l'approche par le niveau dominant ne pouvait montrer. D'après l'EQD (OCDE, 1995), ce n'est que dans cinq pays sur 14 qu'un seul niveau prend plus de la moitié des décisions étudiées. Encore ce niveau peut-il varier. En fait, le cas le plus fréquent, maintenant, parmi les 14 pays, est que deux niveaux dominent ensemble : dans cinq pays, plus de 75 pour cent des décisions sont prises par l'établissement ou par le niveau local.

Enfin, l'EQD a permis de montrer que, pour compliquer encore la situation, certaines décisions pouvaient être prises en collaboration par plusieurs niveaux, autrement dit que d'assez nombreux systèmes éducatifs publics disposent d'une structure intégrée de décisions. Dans huit pays sur les 14, plus de 40 pour cent des décisions sont prises par une collaboration entre deux niveaux de la structure, soit que l'un consulte l'autre, soit que l'un travaille dans un cadre fixé par l'autre (OCDE, 1995). Selon qu'on s'intéresse aux décisions qu'un niveau prend en toute autonomie, à celles qu'il prend éventuellement en collaboration avec un autre, à celles dans lesquelles il intervient, son poids dans la structure peut varier considérablement.

Pour Skilbeck (1990), les raisons de l'insuffisance de l'opposition entre systèmes centralisés et décentralisés sont qu'un seul niveau ne prend jamais toutes les décisions, que l'enseignement privé ajoute à l'hétérogénéité – deux raisons que l'étude quantitative a confirmées –, et enfin que le pouvoir n'appartient pas seulement à celui qui décide, mais aussi à celui qui contrôle ou qui rend compte. L'étude quantitative permet d'ajouter à cela que les décisions peuvent être prises, en dehors du Centre, à des niveaux plus ou moins décentralisés et aussi que la palette des différents modes de décision complique encore l'appréhension de l'importance, même d'un pur point de vue organisationnel, de tel ou tel niveau. Par exemple, il s'avère que lorsque le niveau national prend, dans un pays, un grand nombre de décisions, il a aussi tendance à davantage consulter les autres niveaux avant de les prendre.

Une source de suggestions

A plusieurs égards, les autres études sur le fonctionnement des systèmes éducatifs indiquent comment améliorer l'EQD.

Décision et pouvoir

Ces études rappellent d'abord que les décisions prises par un niveau ne sont pas le seul indice de son pouvoir. Cowen signale par exemple l'influence que le ministère fédéral de l'Éducation exerce, aux États-Unis, en collectant les statistiques relatives à l'éducation et en les diffusant sur l'ensemble du territoire. Pourtant, ces études illustrent aussi les pièges que l'on rencontre dès que l'on veut quitter le terrain solide des compétences telles que les textes juridiques les définissent.

On pourrait donc proposer de compléter l'approche décisionnelle par des indicateurs du type suivant :

- proportion du personnel administratif total exerçant à chaque niveau ;

- proportion des salaires du personnel administratif total versés aux agents administratifs de chaque niveau;
- proportion des dépenses du système éducatif assurées par chaque niveau (« le pourcentage d'argent recueilli à chaque niveau est révélateur de la répartition des responsabilités », voir Holmes, 1979, p. 15).

Décentralisation et démocratisation

Parce que leur approche est plus vaste que celle de la décision, ces études nous mettent aussi en garde contre la tentation d'identifier un système de décision décentralisé avec un système éducatif démocratique. Par système « démocratique », on entend habituellement soit un système favorable à la réussite scolaire des catégories sociales défavorisées, soit un système dont le fonctionnement est soumis à la discussion. Outre qu'il n'existe pas d'évidence que les systèmes décentralisés soient plus favorables aux élèves de catégories sociales défavorisées, il n'est pas sûr non plus que l'intensité et la qualité des débats démocratiques suscités par les systèmes éducatifs soient proportionnelles à leur degré de décentralisation. Les programmes scolaires sont maintenant débattus en France au sein d'un Comité national des programmes, alors qu'ils étaient auparavant décidés souverainement par l'Inspection générale. Les diplômes de l'enseignement technique sont décidés, au niveau national, par concertation entre le ministère et les professions, et pas seulement par le ministère. On peut voir là deux éléments de « démocratisation » qui n'ont rien à voir avec la décentralisation des décisions, même s'il est vrai que certains attendent de la décentralisation une participation plus grande des usagers à la conduite du système. Holmes fait une suggestion intéressante à cet égard : « Il serait utile de faire des études comparatives des instances dans lesquelles, d'un pays à l'autre, les (...) politiques de l'éducation sont débattues » (Holmes, 1979, p. 12). Sans être facile, une mesure directe du degré d'implication de la société civile dans la conduite du système éducatif ne serait pas impossible à réaliser.

D'autres décisions ?

Cependant, l'intérêt le plus grand des études précédentes est que, au fil des comparaisons, y est utilisé, de façon non systématique, tel ou tel critère, qu'on peut confronter à la liste des décisions de l'EQD. On peut alors soit s'interroger sur les décisions considérées dans l'EQD qui ne sont pas évoquées dans ces études, soit se proposer d'ajouter à l'EQD certains des critères qui y sont utilisés.

Les décisions de l'EQD que l'on retrouve rarement dans les études précédentes sont celles qui concernent le plus spécifiquement l'établissement scolaire et ce qui se passe en son sein : création d'établissement, de classe, instance qui détermine l'école fréquentée, mode d'évaluation, organisation de l'aide au travail personnel. Par ailleurs, le financement est le domaine dont le traitement porte le plus la marque de l'approche par l'établissement.

Les décisions ci-dessous sont utilisées dans les études précédentes, ou ont été suggérées par des remarques qui y figurent. Elles pourraient être prises en compte dans

l'EQD, dans la mesure où l'on voudrait élargir davantage cette approche par rapport à celle de l'autonomie de l'établissement scolaire.

a) Décisions sur la formation du personnel :

- durée de la formation des enseignants;
- qualifications requises pour enseigner;
- certification des enseignants, c'est-à-dire quelle instance les déclare aptes à enseigner;
- qualifications requises pour diriger un établissement scolaire;
- organisation de la recherche pédagogique;
- organisation de la formation continue du personnel;
- organisation des services consultatifs fournis aux enseignants.

b) Décisions sur la scolarité des élèves :

- dates de début et de fin de l'année scolaire;
- durée de la scolarité obligatoire;
- organisation des activités extra-scolaires;
- ampleur de la contribution financière des familles à la scolarité;
- montant et critères d'attribution des bourses d'études.

c) Décisions sur l'évaluation :

- organisation du système d'informations;
- définition des indicateurs de compte rendu.

Introduire ces décisions pourrait cependant amener à « décrire les systèmes éducatifs », tâche évidemment sans fin et sans principe directeur. Le fil directeur de l'EQD a été jusqu'ici l'intérêt porté par les responsables des systèmes éducatifs à l'autonomie de l'établissement. L'insertion de ce travail dans les indicateurs de l'enseignement peut permettre de trouver un autre fil directeur, qui articulerait l'approche des décisions à tel ou tel pan du système des indicateurs.

Enfin, ces études ont permis de vérifier qu'établir la liste des définitions suppose aussi des choix sur leur caractère plus ou moins synthétique. Par exemple, à lire un ouvrage sur les examens (Eckstein et Noah, 1993), on voit bien que, dans l'EQD, sous la décision S6 (Construction des examens), sont rassemblées un nombre considérable de décisions concernant le choix des sujets, celui des critères de jugement, des correcteurs, de la durée des épreuves, etc. La définition d'une décision intègre donc un jugement sur son importance par rapport aux autres, et sur le caractère lié des « sous-décisions » qu'elle contient (lié signifiant en l'occurrence prises plutôt par la même instance et selon le même mode).

3. Comparaison avec d'autres caractéristiques des systèmes éducatifs

Une première phase du travail sur les processus de décision s'achève. Avant d'entamer la deuxième, il est important de pousser le plus loin possible l'utilisation des premiers résultats, quitte à prendre des risques méthodologiques. En particulier, il est important d'utiliser le fait que les indicateurs des processus de décision font partie d'un ensemble d'indicateurs qui couvrent de multiples autres aspects des systèmes éducatifs.

En l'état actuel des travaux, il ne s'agit pas de produire des résultats infaillibles. Plus modestement, on se limite à dégager les voies possibles d'une approche future. On se contente d'articuler quelques indicateurs. Cette approche comporte des inconvénients, parce que les différents indicateurs ne couvrent pas toujours les mêmes pays, les mêmes niveaux et types d'enseignement. Pour ces raisons, le choix des indicateurs est parfois dicté par la disponibilité des données plutôt que par leur pertinence.

Le processus de décision est ici rapproché de trois autres caractéristiques des systèmes éducatifs figurant dans la publication *Regards sur l'éducation* (OCDE, 1992 et 1993) : la réussite à la fin du second cycle (R5), les dépenses publiques et privées par élève (P6), et le niveau de lecture des élèves de 14 ans (R1). (Les valeurs des indicateurs utilisés ici pour décrire les processus de décision sont tirées de la publication de l'OCDE parue en 1995.) Pour des raisons qui sont fournies dans ce rapport, elles peuvent être parfois légèrement différentes de celles qui sont données dans *Regards sur l'éducation*.

La réussite à la fin du second cycle

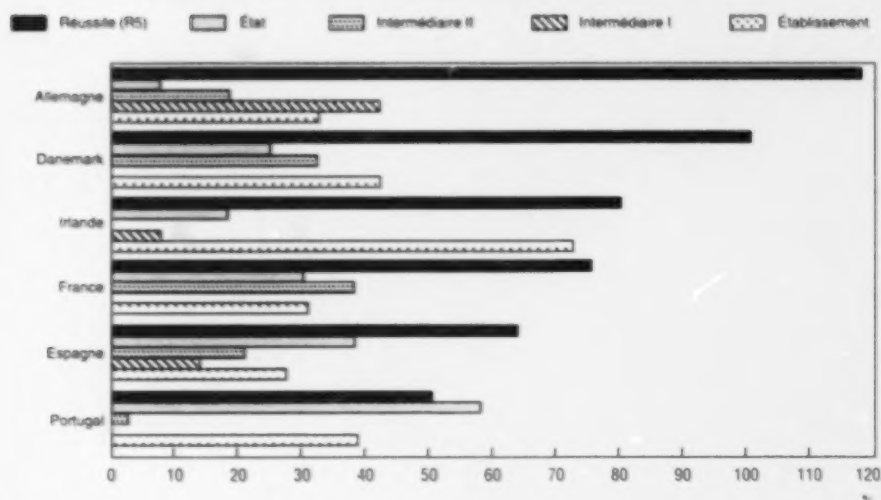
Le graphique 2.1 illustre le croisement des indicateurs « Proportion de décisions prises par les différents niveaux pour l'enseignement de degré CITE 3 » et « Réussite à la fin du second cycle du secondaire » (R5). Ce dernier mesure la proportion d'une génération qui obtient un diplôme de fin d'études secondaires. On lit par exemple dans ce graphique qu'au Danemark, 100 pour cent des élèves d'une génération obtiennent un diplôme de fin d'études secondaires, le niveau national prend un peu plus de 20 pour cent des décisions, etc. L'indicateur des niveaux de décision se limite à l'enseignement public, alors que le taux de réussite porte aussi sur l'enseignement privé. Ceci ne devrait pas trop biaiser les comparaisons puisque, dans tous les pays retenus ci-dessous, l'enseignement public est largement majoritaire.

Dans les deux pays où les taux de réussite sont les plus faibles, c'est-à-dire au Portugal et en Espagne, le niveau national (État) prend plus de décisions que les autres niveaux. À l'inverse, dans les deux pays où les taux de réussite sont les plus élevés, en Allemagne et au Danemark, l'État en prend moins que les autres niveaux.

Peut-on en déduire que le poids de l'État dans les décisions est inversement proportionnel à la possibilité pour les élèves de terminer avec succès leurs études secondaires ? Est-ce une vérification de l'hypothèse évoquée par Haag (1981, p. 92) selon laquelle la décentralisation « faciliterait l'extension du droit à l'éducation, parce qu'elle permettrait de mieux appréhender les réalités locales » ?

Évitons les conclusions hâtives. L'indicateur R5 n'est pas univoque puisqu'il témoigne à la fois de la difficulté des examens, du degré d'ouverture du système éducatif et de son efficacité. Lequel de ces trois phénomènes est lié à la décentralisation du système ?

Graphique 2.1. Réussite à la fin de CITE 3 et niveaux de décision
Pass rate at the end of ISCED 3 and levels of decision-making



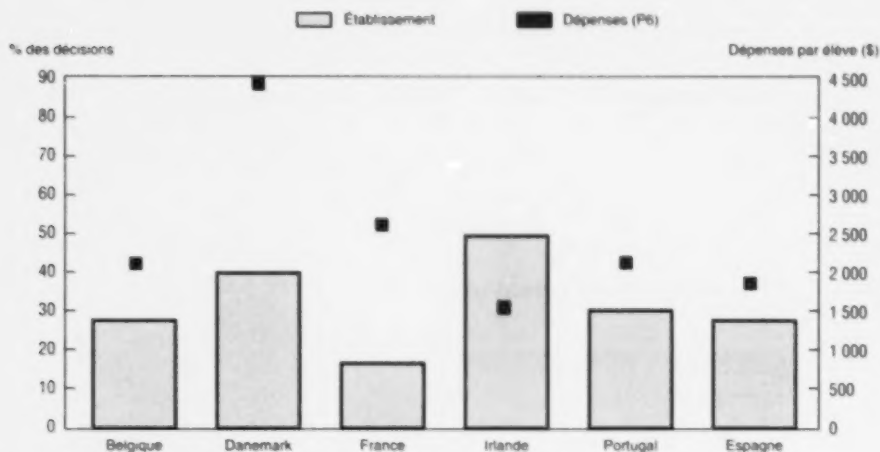
Source : A partir de données OCDE.

D'autre part, R5 n'est pas dépourvu d'ambiguïtés techniques. Dans les pays où les taux de réussite sont élevés, ils peuvent dépasser les 100 pour cent. Est-ce que cela s'explique uniquement par le fait qu'une partie des diplômés est plus âgée que la population de référence, ou est-ce que les données sont saisies différemment ? Enfin, la réussite à la fin du secondaire dépend de l'ensemble de la scolarité, et pas seulement de la scolarité en CITE 3. Dans les pays qui figurent sur le graphique, à l'exception du Danemark, le processus de décision est identique, ou presque, en CITE 2 et 3. Si ce n'était pas le cas, il faudrait décider comment combiner la mesure des processus de décision dans ces deux niveaux.

Les dépenses par élève

Le graphique 2.2 met en rapport les dépenses publiques et privées par élève et la proportion des décisions prises par l'établissement, pour l'enseignement primaire public. On y lit par exemple que les écoles primaires danoises prennent 40 pour cent des décisions qui les concernent, tandis que l'on dépense dans ce pays environ 4 500 dollars sur fonds publics ou sur fonds privés, par élève de l'enseignement primaire, public ou privé.

Graphique 2.2. Dépenses par élève et décisions prises par l'établissement, enseignement primaire public
Expenditure per student and decision-making by school, public primary education



Source : A partir de données OCDE.

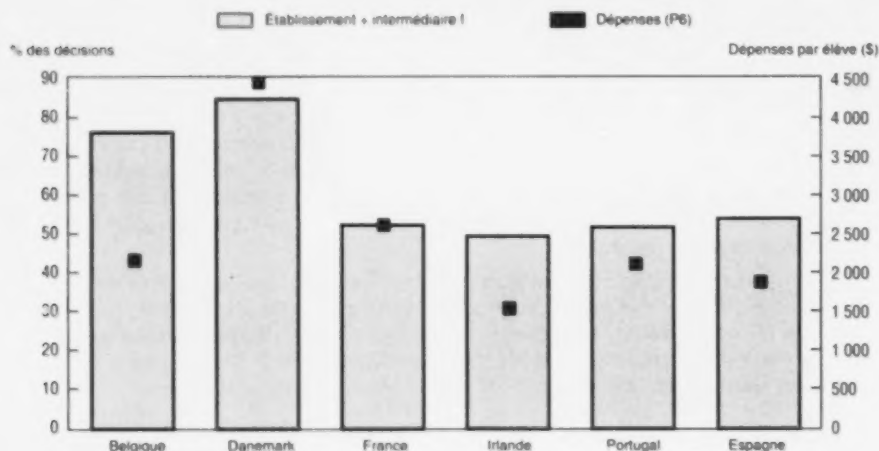
Les différences entre pays semblent dues au hasard. Avec une part faible ou forte de l'établissement dans la prise de décision, les dépenses par élève peuvent être fortes ou faibles.

On sait que les communes supportent souvent en grande partie les coûts de l'enseignement primaire. Dans le graphique 2.3, les décisions prises au niveau de l'établissement et au niveau le plus proche de l'établissement (qui ne comporte pas exclusivement les communes) sont agrégées.

L'image obtenue semble plus cohérente que celle du graphique 2.3. Excepté la Belgique, les dépenses par élève semblent liées à la part des décisions prises aux deux niveaux considérés. Si, comme au Danemark, beaucoup de décisions sont prises de façon relativement proche de l'élève (établissement et intermédiaire inférieur), l'enseignement semble coûter plus cher, soit que les coûts de gestion du système soient plus élevés, soit qu'on ait alors plus de difficultés à fermer certaines écoles coûteuses. Pour faire des économies, suffirait-il alors de centraliser les systèmes éducatifs ?

Évidemment, le nombre de pays n'est pas représentatif et l'on n'a pas cherché d'autres facteurs explicatifs, il vaut donc mieux s'abstenir de conclure. D'autre part, les dépenses privées, qui figurent dans l'indicateur, sont peut-être indépendantes de la structure de décision et peuvent accroître indûment les dépenses de certains pays.

Graphique 2.3. Dépenses par élève et décisions prises par l'établissement
et le niveau intermédiaire I
Expenditure per student and decision-making by school and intermediate level I



Source : A partir de données OCDE.

Le niveau de lecture des élèves de 14 ans

On trouvera en annexe six graphiques (2.4 à 2.9) qui illustrent les liens entre le niveau de lecture des élèves de 14 ans, tel que mesuré dans *Regards sur l'éducation* (OCDE, 1993, indicateur R1) à partir de l'étude de l'IEA « Reading literacy », et l'importance des différents niveaux de décision dans un système éducatif.

Ce deuxième aspect est successivement mesuré par la proportion de décisions prises pour le degré CITE 2 public, le niveau national, par le niveau régional (I supérieur), par le niveau local (I inférieur), et enfin par l'établissement lui-même. Il semble qu'il y ait une absence totale de liens entre la structure de décision d'un système éducatif et le niveau de ses élèves en lecture. Le graphique 2.9 pourrait laisser penser que les systèmes où de nombreuses décisions se prennent soit au niveau local soit dans l'établissement ont peut-être des performances un peu plus élevées en lecture. En réalité, le coefficient de corrélation est quasiment nul (0,08) entre ces deux séries de données.

Bien évidemment, la comparaison de la structure de décision avec les connaissances acquises par les élèves devrait se faire sur un panier de plusieurs disciplines et sur plusieurs observations. D'autre part, il faudrait peut-être rapprocher la structure de décision en CITE 2 de disciplines dont l'apprentissage a aussi commencé en CITE 2. Il

faudrait aussi mesurer le niveau de lecture seulement pour le public. Ce résultat n'a donc qu'une valeur extrêmement limitée.

Les exemples ci-dessus sont intéressants dans la mesure où ils soulèvent des questions qui donnent envie d'approfondir le sujet. En outre, ils mettent à nu les problèmes qui empêchent de faire des comparaisons significatives.

Avant tout, les différents indicateurs devraient couvrir la même chose. Ceci concerne surtout les pays pris en compte et les niveaux CITE. Par exemple, des comparaisons entre cinq ou six pays ne permettent guère de savoir si l'on a affaire à des cas typiques ou uniques. De même, il est fâcheux de tomber sur un indicateur qui ne distingue pas entre niveaux CITE 2 et 3, le tout étant classé dans «secondaire» : les enseignements obligatoire et post-obligatoire se trouvent ainsi mélangés. Enfin, il importe que la caractéristique du processus que l'on veut évaluer soit mesurée au long de la période de formation du résultat.

Les systèmes d'enseignement étant complexes, leur analyse le sera également. L'articulation de deux indicateurs seulement ne donnera que des explications partielles et peut créer des malentendus. Une analyse multi-variée – les différents indicateurs constituent les variables – peut être envisagée. Une autre voie serait de désagréger les différents indicateurs avant de les articuler en fonction d'objectifs clairement précisés.

4. Conclusion

Les rapprochements ci-dessus ne plaident pas pour l'existence d'un lien entre la structure des systèmes éducatifs et leur efficacité. Si l'on ajoute à cela ce qui a été montré dans le début du chapitre, à savoir le rôle qu'a joué un certain basculement des évidences dans la réforme des structures des systèmes éducatifs, il serait tentant de conclure que ce basculement a peut-être distrait l'attention de moyens plus adéquats pour augmenter l'efficacité.

D'une part, ces résultats doivent être pris avec toutes les précautions requises par les limites méthodologiques qui ont été indiquées, d'autre part ils ne portent que sur une partie des indicateurs d'efficacité. Par exemple, les tenants de la décentralisation, en France, en attendent plutôt des effets sur l'insertion professionnelle que sur l'acquisition des compétences de base. Il se peut aussi que les réformes qui ont contribué à accroître le pouvoir des échelons les plus bas du système visaient d'autres objectifs que l'efficacité. Il se peut aussi que des études longitudinales aboutissent à la conclusion que, dans un pays donné, une centralisation ou une décentralisation des décisions a été positive, sans pour autant changer sa place par rapport aux autres.

D'un point de vue méthodologique, il apparaît que la mise en relation d'indicateurs de nature différente requiert l'élaboration préalable de modèles de type causal, même rudimentaires. En particulier, une grande attention doit être portée à l'articulation des populations, des périodes et des parties (publique et/ou privée) du système pour lesquelles on mesure d'une part les processus et d'autre part l'efficacité.

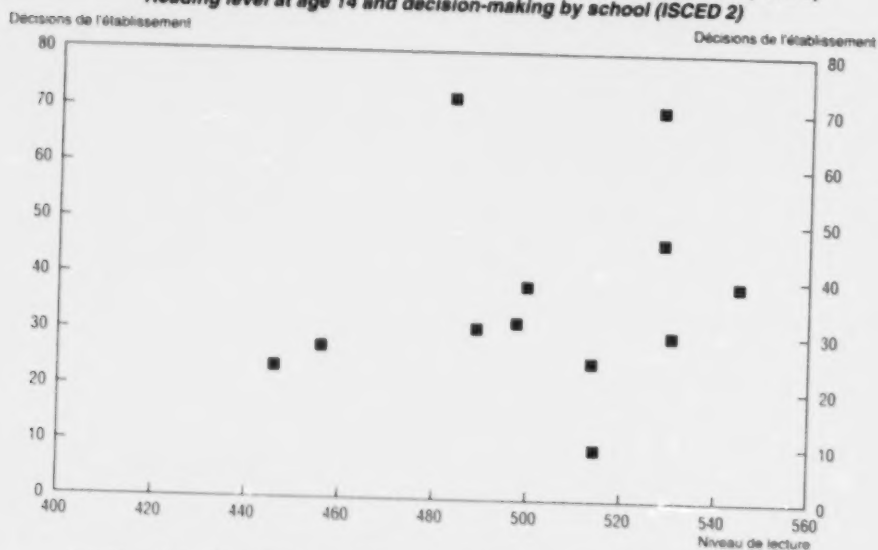
Références

- COWEN, R. (1982), *Les structures de l'éducation*, Annuaire international de l'éducation, Vol. XXXIV, UNESCO, BIE, Genève.
- CROZIER, M. (1963), *Le phénomène bureaucratique*, Seuil, Paris.
- ECKSTEIN, M.A. et NOAH, H.J. (1993), *Secondary School Examination : International Perspectives on Policies and Practice*, Yale University Press, New Haven.
- Eurydice (1991), *Répartition des compétences dans le domaine de l'éducation, situation dans les douze états Membres de la Communauté européenne*, publié pour la Commission des communautés européennes, Bruxelles.
- HAAG, D. (1981), *Pour le droit à l'éducation, quelle gestion?*, UNESCO, BIE, Genève.
- HEATH, K.G. (1962), *Ministries of Education, their Functions and Organisation*, US Department of Education and Welfare, Washington, DC.
- HOLMES, B. (1979) (sous la dir. de), *Guide international des systèmes d'éducation*, UNESCO, BIE, Genève.
- KING, A. (1967), *Organisation et buts de l'enseignement*, Rapport présenté à la conférence internationale sur la crise mondiale de l'éducation, Williamsburg, Virginie.
- OCDE (1993), *Regards sur l'éducation : les indicateurs de l'OCDE* (bilingue), CERI, Paris.
- OCDE (1995), *Les processus de décision dans quatorze systèmes éducatifs de l'OCDE*, CERI, Paris.
- RANSON, S. (1985), "Changing relations between centre and locality in education", in MacNay, I. et Ozga, J. (éd.), *Policy Making in Education. The Breakdown of Consensus*, Pergamon Press / Open University Press, Oxford.
- SKILBECK, M. (1990), *La réforme des programmes scolaires, où en sommes-nous?*, OCDE, Paris.

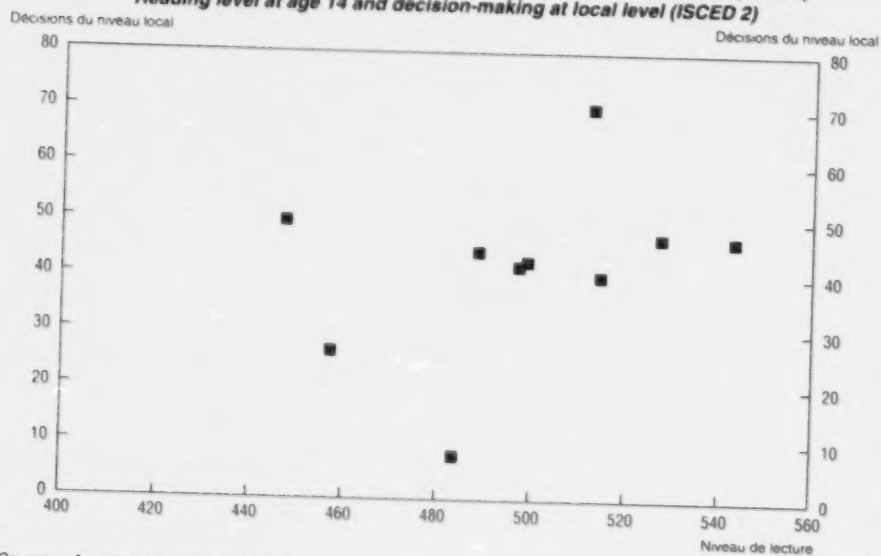
Annexe

**Relations entre le niveau de lecture à 14 ans
et le poids des différents niveaux administratifs dans la structure de décision**

Graphique 2.4. Niveau de lecture à 14 ans et décisions de l'établissement (CITE 2)
Reading level at age 14 and decision-making by school (ISCED 2)

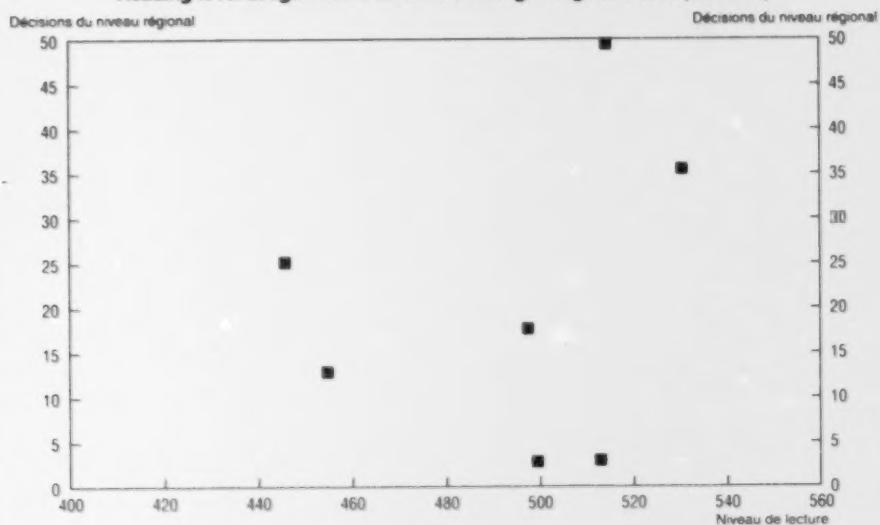


Graphique 2.5. Niveau de lecture à 14 ans et décisions du niveau local (CITE 2)
Reading level at age 14 and decision-making at local level (ISCED 2)

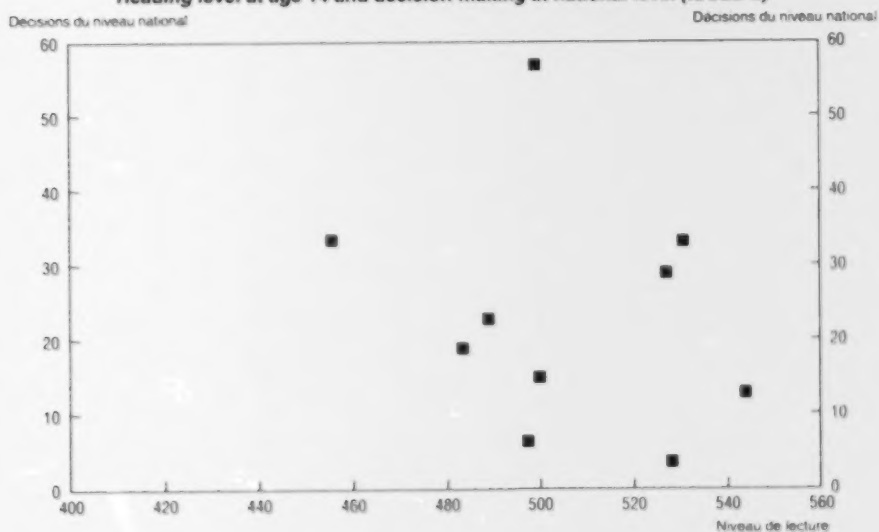


Source : A partir de données OCDE

Graphique 2.6. Niveau de lecture à 14 ans et décisions du niveau régional (CITE 2)
Reading level at age 14 and decision-making at regional level (ISCED 2)

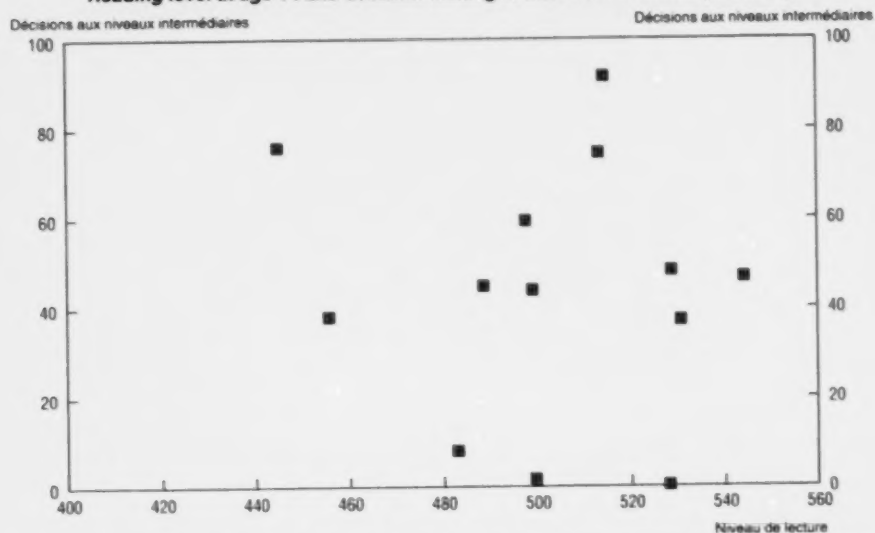


Graphique 2.7. Niveau de lecture à 14 ans et décisions du niveau national (CITE 2)
Reading level at age 14 and decision-making at national level (ISCED 2)

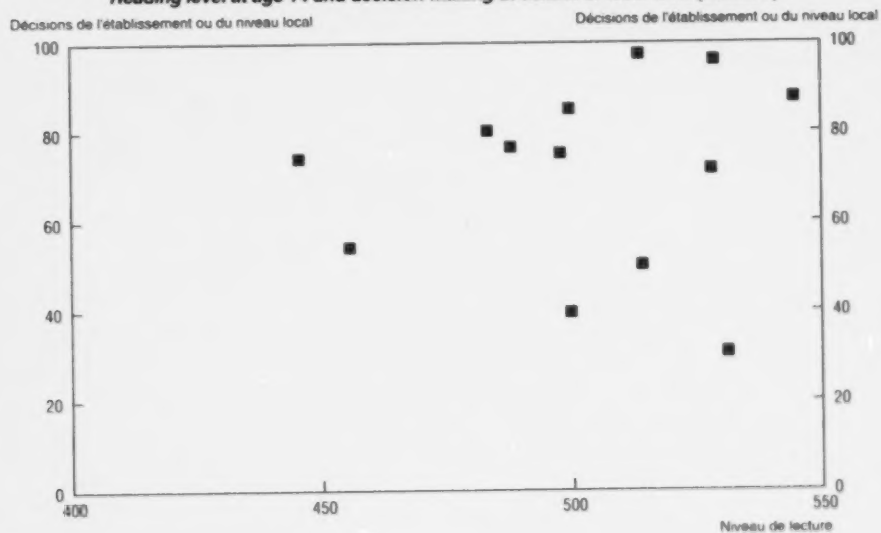


Source : A partir de données OCDE.

Graphique 2.8. Niveau de lecture à 14 ans et décisions aux niveaux intermédiaires (CITE 2)
Reading level at age 14 and decision-making at intermediate levels (ISCED 2)



Graphique 2.9. Niveau de lecture à 14 ans et décisions de l'établissement ou du niveau local (CITE 2)
Reading level at age 14 and decision-making at school or local level (ISCED 2)



Source : A partir de données OCDE.

The Selection and Definition of International Indicators on Teachers
*Sélection et définition des indicateurs internationaux
sur les enseignants*

by

Jaap Scheerens

Department of Education, University of Twente, the Netherlands

In this chapter the relevance and feasibility of teacher indicators are discussed. Three rationales are described, from which teacher indicators can be generated: general macro-level description, effectiveness and efficiency perspective and the political economy of the teaching profession. This chapter also reveals the problems and likely pitfalls of international comparisons. The chapter reviews teacher compensation, the working time of teachers, teacher qualifications and stability of the national teaching force. Standard definitions and a structured questionnaire for data collection are described. For some indicators, notably stability of the national teaching force, and type of teacher training, data availability was so low that indicators on these topics had to be abandoned.

Some of the interesting results on the remaining four indicators (teaching time, duration of teachers' education, teacher compensation and teacher characteristics) are the following:

- Both at the primary and secondary level (ISCED 1 and 2) there is considerable variance in teaching time between countries. This variation is even higher at ISCED level 3.
- The duration of teachers' education is rather uniform: in most countries teachers have had 15 or 16 years of education before they are allowed to teach at ISCED levels 0, 1 and 2.
- There are fairly large differences in salaries for teachers in OECD countries, both in levels of pay and structures of pay scales.
- In nearly all countries, the teachers aged from 40 to 49 form the largest group in public primary and secondary education. A large majority of teachers at ISCED levels 0 and 1 are female. At ISCED level 3, men and women appear to be equally present.

These results are discussed further with respect to trustworthiness and relevance to policy-making. Some suggestions for improving the available databases on teachers are provided.



Note de synthèse

On étudie dans ce chapitre la pertinence et l'adéquation des indicateurs relatifs aux enseignants.

On décrit trois modalités d'élaboration des indicateurs sur les enseignants : la description générale, de niveau national, « l'optique de l'efficacité et de la rentabilité », et une autre optique intitulée « économie politique du corps enseignant ». Sont inclus dans cette dernière catégorie les indicateurs de l'offre et de la demande d'enseignants, les aspects financiers de base, la charge de travail et les possibilités de carrière.

En examinant les études internationales comparatives conduites sur les enseignants, on constate l'existence de variations considérables entre pays quant aux conditions de travail des enseignants. Cet examen met aussi en lumière les problèmes et les obstacles probables de ces comparaisons internationales.

On se sert des enseignements tirés de ces études, et des résultats d'une enquête sur la disponibilité de données relatives aux enseignants dans les pays de l'OCDE, pour préciser et affiner encore les indicateurs de la rémunération des enseignants, de leur temps de travail, de leurs qualifications et de la stabilité du corps enseignant national. Les définitions normalisées, ainsi qu'un questionnaire structuré en vue de la collecte des données sont décrits dans ce chapitre.

Les données ont été obtenues par l'intermédiaire des correspondants nationaux à partir des sources disponibles, à savoir, les statistiques nationales, les documents officiels ou les enquêtes représentatives.

Pour certains indicateurs, notamment ceux qui concernent la stabilité du corps enseignant et le type de formation pédagogique, la disponibilité des données semble si réduite que l'on a dû renoncer à présenter des indicateurs sur ces aspects. On trouvera ci-dessous quelques-uns des résultats les plus intéressants obtenus à propos des quatre indicateurs restants (temps d'enseignement, durée de la formation des enseignants, rémunération et caractéristiques des enseignants) :

- *tant au niveau primaire que secondaire (CITE 1 et 2), le temps d'enseignement varie considérablement d'un pays à l'autre. Cette variation est plus forte encore au niveau CITE 3 (enseignement professionnel), allant de 612 heures par an (Italie, Portugal et Suède) à 1 692 (Turquie) ;*

- pour ce qui est de la durée de la formation des enseignants, on constate une certaine uniformité : dans la plupart des pays, les enseignants font quinze ans d'études avant d'être autorisés à enseigner aux niveaux CITE 0, 1 et 2. C'est l'Allemagne qui exige le plus grand nombre d'années de formation pédagogique pour tous les niveaux CITE;
- il existe d'assez fortes différences de salaires entre les enseignants des pays de l'OCDE. On constate aussi de sensibles différences entre le nombre d'années nécessaires pour obtenir le salaire maximal (huit ans au niveau CITE 2 en Nouvelle-Zélande, contre 45 ans en Espagne et en Italie). Si on les compare au niveau général de vie d'un pays, les salaires des enseignants débutants dans des pays tels que l'Espagne et la Turquie sont assez élevés, ce qui pourrait dénoter le prestige dont y jouit le métier d'enseignant;
- dans la quasi-totalité des pays, la tranche d'âge de 40 à 49 ans est la plus nombreuse dans les enseignements primaire et secondaire publics. On trouve une forte majorité de femmes parmi les enseignants des niveaux CITE 0 et 1. Au niveau CITE 3, hommes et femmes semblent être à égalité.

Ces résultats sont étudiés de plus près en tenant compte de leur fiabilité et de leur pertinence politique. Quelques suggestions sont présentées en vue d'une amélioration des bases de données sur les enseignants.

*
* *

1. Introduction

Roughly, the INES conceptual framework has the structure of an *input, process, output and context systems model*. While the validity and usefulness of this linear model is still debated, the work on measuring some of its aspects proceeds. Indicators on teachers fall partly in the input (or resources) category and partly in the process category. Features such as pupil : teacher ratio, levels of training of teachers, years of experience and shortages of teachers are classified in the input category, whereas indicators that could reflect the relative effectiveness in professional practice or technique belong to the category of process indicators.

It should be noted that the importance of indicators on teachers can hardly be overemphasised, both from a financial perspective, as the budget for teachers' salaries is by far the most important part of national education budgets, and from a substantive educational perspective, as the success of all educational policy will ultimately depend on the implementation by the teaching force. It may be added, however, that narrowly descriptive statistics on teachers have less value in an OECD context.

During the initial phases of the INES project, only two indicators were developed and calculated in the area of teachers, namely "staff employed in education" and "pupil : teacher ratio". During the third phase of the project, more teacher indicators have been developed. So far, the work of Network C in the area of developing additional

teacher indicators has concentrated on four themes: teacher working time, teacher compensation (salaries, in particular), teacher qualifications and the age and sex distribution of the teaching population in a country. The conceptual development of indicators in these domains will be summarised in a subsequent section. In preparation for measuring these indicators, several issues are in need of further clarification.

First of all, in an effort to motivate the incorporation of particular indicators on teachers, the rationales for the selection of these will be made explicit. Second, it is worth summarising what is already known about international comparisons on teachers. Third, the available studies clearly illustrate the difficulties in making international comparisons on teachers; these difficulties, as well as certain limitations of current international statistics, will have to be made explicit and faced in the operational procedures that will be proposed for the INES indicators on teachers. Fourth, reference will be made to standard definitions, which are needed for key concepts in the area of teacher indicators, namely the distinction between ordinary teachers and other non-teaching professionals who are employed in the field of education; the standard (ISCED) qualification of levels within the education system, the computation of full-time equivalents; the distinction between basic activities, such as actual classroom teaching, other duties that are part of the formal job of a teacher, and "extra" activities of teachers; and the type of salary data (gross or net) that is to be used.

Next, the actual items (*i.e.* questions) that are used to gather the basic data elements for the teacher indicators will be specified. Finally, the results of a first wave of data collection that took place in January and February 1994 will be summarised.

2. Rationales for Considering Teacher Indicators

Like no other service or "business", education is dependent on human resources. This in itself could be taken as sufficient cause to invest efforts in the development of internationally comparative indicators on teachers. In order to focus the attention further and to come to grips with the question of which specific aspects of the teacher domain should be covered, certain perspectives will be used to analyse this field.

The first perspective is a global review of the issues that presently confront the teaching profession. The other three perspectives to be considered are more closely related to the question of which topics are to be considered as likely candidates for indicators. The perspectives are "macro-level description", "effectiveness and efficiency" and "the political economy of the teaching profession".

Issues that Presently Confront the Teaching Profession

In publications such as *The Teacher Today* (OECD, 1990) and *The Teaching Nation* (Neave, 1992), some important trends and problem areas concerning the teaching profession are discussed. In summary, these amount to increased demands and expectations of the teaching force on the one hand, and limitations with respect to compensation, both monetary and other, on the other. In addition, demographic developments concerning both student populations and teaching forces are sometimes taken as negatively affecting

the career expectations of teachers. The inherent tensions are seen as the basis of what Neave (1992) calls a malaise in the teaching profession.

One of the increased demands on the teaching force is the tendency to delegate to schools a growing part of the informal upbringing of children and adolescents in such domains as the affective and the social. Other requirements mostly originate from national policy initiatives such as the use of modern information technology in school, the integration of special and "regular" education, and more individualised tuition. In some countries at least, there is the general impression that, partly because of the liberal pedagogical tendencies of parents during the 1970s, children are more difficult at school and provide more disciplinary problems. The influx of cultural minorities in the schools of most Western European countries, and the integration of these pupils in the school system, can also be seen as a factor that increases the difficulty of the teacher's job. Of course, it would be drawing a too one-sided picture if each and every aspect of modernisation of teaching were simply interpreted as an increased burden. It certainly does apply, however, to the increased autonomy given to teachers and the renewed call for professionalism in teaching, as in the American *restructuring* movement. There are new demands implicit in these, such as investing more time in participatory decision-making on the school's policy and in contacts with parents and other constituencies to which the school is expected to be responsive. But on the other hand, this philosophy of "teacher empowerment" can be seen as a part of "re-valuation" of the teaching profession and of enhancement of teachers' motivation.

A similar kind of reasoning applies to increased accountability demands that often accompany a policy of devolution of authority and deregulation of educational processes. Although the accountability issue is quite complicated, both in a technical and political sense, assessment and feedback within the framework of accountability requirements could also be seen as stimulating responsiveness and organisational learning. In actual practice, resistance to accountability requirements, however, is likely to have the upper hand and, from the point of view of teachers, they will be added to the overall impression of increased demands.

Demographic tendencies, such as a dramatic decline in pupil enrolments in most Western European countries after 1975, policies of contraction of education budgets during the 1980s and a "spirit of the times" that is more inclined to value private enterprise than public service can be seen as background factors to the fact, or rather the subjective impression among teachers, that the increased job demands described above have been less than adequately compensated for.

The decline in birth rates has lead to an over-supply of teachers, to unemployment and to an increased proportion of the teaching population having "precarious employment prospects" (Neave, 1992, p. xiii), for instance, temporary posts and part-time appointments. In specific areas, notably the exact sciences and technical and vocational training, shortages have arisen, which could be attributed to the stronger competitive position of the private sector as an employer compared to education. Over-supply of teachers is taken as a cause of decreased teacher mobility and of a blocking of career prospects. Another related consequence of these demographic tendencies is a greying of the teacher population in the countries where they have occurred.

Teacher compensation is not only a matter of salaries, but also of public recognition and valuing of the teaching profession. Ultimately, the balance of job demands and various kinds of compensation, among which the intrinsic motivation to teach should not be forgotten, will be the basis for the job satisfaction of teachers.

The objective data that are available on teachers' salaries and public estimation of the teaching profession do not altogether explain the subjective impression of a malaise surrounding the profession. The summary tables in which data on teachers' salaries are compared to other manual and non-manual jobs generally indicate that teachers earn more than comparable groups and show lower earnings only among those first entering the profession (see, for instance, the tables quoted from the European Commission's *Final Report on the Conditions of Employment* in Neave, 1992, pp. 97 and 99). A recent Dutch study (De Jonge *et al.*, 1992) indicates that in 1979 teachers still had an advantage over civil servants of three percent in salary, whereas in 1991 teachers on average fell 10 per cent behind comparable civil servants. Another Dutch study (Leune, 1988) indicates that the social status of teaching has risen in the Netherlands since the end of the Second World War. Nevertheless, the overall attractiveness of becoming a teacher has gone down (Ministerie van Onderwijs en Wetenschappen, 1989).

These current issues surrounding the teaching profession underline the relevance of indicators in the general areas of teacher supply and demand (including shortages in certain subjects), unemployment, working conditions, salaries and valuation of the teaching profession by teachers themselves and by the general public.

Macro-level Description

Although the relationships between the input, process and output indicators of the conceptual framework of the INES project should be seen as loosely coupled, nevertheless the most important rationale for selecting input and process indicators is that these may provide information that could explain variance in output and outcome indicators (where outputs are taken as the more immediate results of schooling such as achievement in certain subjects and attainment, *i.e.* completion and graduation; and outcomes are understood as the ultimate societal effects of education such as employment and earnings). In this, there are differences to a degree, however. The most direct relationship between inputs and processes on the one hand, and outputs on the other, is implied when using the research literature on educational productivity and school effectiveness to select the input indicators (*cf.* Scheerens, 1990). This is the perspective that will be used in the next sub-section. It could also be maintained, however, that there is a category of teacher indicators that is still to be considered relevant even though the relationship with outputs and outcomes is quite remote. Descriptive statistics on the teaching force of a country can be considered relevant from the perspective of providing estimates of the magnitude of the investments of countries in human resources for education, in other than strictly monetary terms. One example is the indicator on the total (teaching and non-teaching) staff in education that has already been calculated within the INES project. This particular indicator can be seen as the counterpart of financial indicators on educational budgets and expenditure per pupil. Other examples of prospective macro-indicators that might be chosen from this perspective are statistics on the age and sex distribution of the teachers

of a certain school level in a country and perhaps also indicators on over-supply and shortages of teachers in certain subjects, although these may well have a more direct influence on educational outcomes.

Effectiveness and Efficiency

"Educational effectiveness" refers to the degree to which educational inputs or processes can be shown to affect educational goal attainment; in other words, to influence whether these antecedent conditions have a positive relationship with desired outputs or outcomes. When, in addition, the costs of the inputs and processes can be accounted for, conclusions about efficiency can also be drawn. Translated to the problem at hand, *i.e.* the selection of teacher indicators, the effectiveness/efficiency perspective means looking for those teacher variables that have been shown empirically to have a significant positive association with educational attainment.

It should be noted that, in itself, the effectiveness/efficiency concept is not restricted to a specific domain of education goals. The results of any search for variables that have been investigated in empirical research, however, will be limited to the particular type of educational attainment that has been assessed in the studies concerned. This means that the knowledge we have on the effectiveness of educational conditions, teachers included, is particularly relevant in the area of cognitive achievement in basic school subjects, such as reading, mathematics, science and proficiency in the mother tongue. (There is, however, an increasing number of studies in which non-cognitive outcomes such as interest in subjects and self-image of students are included.)

There are in fact two distinct traditions in educational effectiveness research that are relevant for the selection of teacher variables. The first is known as *input-output* studies and has an economic and sociological disciplinary background, and the second is known as *process-output* studies and has its disciplinary basis in educational psychology.

Within the framework of the input-output studies, several teacher-related variables have been the object of attention in many studies, namely: teacher : pupil ratio, teacher qualifications or preparedness, teacher experience, teachers' salaries and per pupil expenditure. The results of the studies in which these variables were related to achievement measures have been rather disappointing. In fact, many reviews and meta-analyses have indicated that these variables have not been found to be positively associated with achievement in any significant and consistent way (*cf.* Hanushek, 1986). In the case of teacher experience, for example, Hanushek looked at 109 studies; of these, 33 indicated a positive and significant association with achievement, in seven studies a significant negative association showed up, and in 69 studies a non-significant relationship was found, with almost as many negative as positive signs. For the other above-mentioned variables, the results were even less significant and consistent. The explanation of these counter-intuitive results could be that in industrialised countries conditions are such that there is not enough variance in these antecedent conditions to affect outcomes in a substantial way.

At the same time, researchers go on trying to obtain better data and more sensitive research designs, so that small but nevertheless socially meaningful effects might still be detected. For example, in the area of teacher preparedness several recent studies have

indicated that preparedness, in the sense of knowledge of the subject that they teach, but also in the sense of pedagogical knowledge, matters as far as achievement is concerned (Ferguson, 1991; Hanushek, 1986; Monk, 1993). Also, Hedges *et al.* (1994), after a re-analysis of Hanushek's meta-analyses, contest his results and arrive at somewhat more positive conclusions about the impact of per pupil expenditure.

What conclusion should we draw from this rather disturbing research evidence? Although one should be sensitive to the fact that research is never perfect, the evidence is quite impressive and should not be simply disregarded. These counter-intuitive results are informative, since they make it clear that one should look at other areas for explanations of "what works in education". Next, the use of the effectiveness/efficiency perspective as the sole rationale to select indicators, although quite sensible, should be looked upon with a certain degree of relativity; in an earlier section the loosely coupled inter-relationship between the input, process and output indicator categories of the conceptual framework of the INES project has been referred to. Perhaps a compromise solution could be to put some effort into constructing internationally comparable indicators, such as those referring to salaries, experience, pupil : teacher ratio and teacher qualifications, but at the same time to make a major effort to prepare more process-oriented teacher indicators. In this respect, the second research tradition that we are about to discuss, the process-output studies, has indeed provided more consistent results.

Process-product research is an important stream in the area of research on teaching. It focuses particularly on classroom processes, instructional strategies and classroom organisation by the teacher. The results of these studies have been summarised in reviews and research syntheses (for example, Brophy and Good, 1986; Walberg, 1986; Fraser *et al.*, 1987; Wang *et al.*, 1993).

The results, which have also been supported by evaluations of more holistic instructional approaches such as *mastery learning* and *direct teaching* (cf. Smith and O'Day, 1988; Scheerens, 1992; Creemers, 1994), underline three basic factors: optimisation of instructional time, opportunity to learn, (*i.e.* correspondence between content taught and content assessed) and structure. This last factor, structured teaching, emphasises clear instructional objectives, frequent review and questions, feedback and correctives.

In previous phases of the INES project, indicators on "time on task" and "content covered" have been conceptualised. As the calculation of these indicators requires data at classroom level, for which the INES project appears to be dependent on available international survey data, they will not be further considered in this chapter. Nevertheless, the importance of such indicators on actual teacher functioning should again be stressed.

The Political Economy of the Teaching Profession: System-Level Conditions that Affect Teachers' Motivation

In organisational theory, the productivity concept of organisational effectiveness, which is in fact the effectiveness concept implied in educational effectiveness research, is seen as just one of several effectiveness concepts. Alternative effectiveness criteria are an organisation's continuity, its adaptiveness or responsiveness to environmental developments, and the motivation and satisfaction of the members of the organisation. Even if the

position is taken that these alternative criteria of organisational effectiveness are to be considered as means to the ultimate (productivity) criterion (Scheerens, 1992), such alternative effectiveness criteria may be used as an additional perspective to select indicators on teachers. As stated in the section on issues that presently confront the teaching profession, much of the current debate on the position of teachers can be seen as an analysis of costs and benefits of being a teacher. The central issue appears to be the desirability of a well-motivated, flexible teaching force that is in a good position to respond to the changing and increased demands of societies.

The more or less objective conditions that form the basis of the attractiveness of the teaching profession in general, and of specific school levels and subject-matter areas in particular, are only part of the story. As Neave (1992, p. 3) points out, although operational indicators such as salary, scales of promotion and working hours may be important, the larger context of historical developments in the valuation of the teaching profession, demographic developments and the spirit of the time also count. Nevertheless, we shall have to limit ourselves to these more or less objective conditions that are relevant to motivational aspects of the teaching force, and which in their turn may influence the professional self-image of teachers as well as the status of teachers perceived by the society.

There are four major categories of system-level conditions that will affect the attractiveness of being a teacher:

- supply and demand of teachers,
- basic financial conditions,
- work-load, and
- career opportunities.

Supply and Demand of Teachers

As pointed out before, in certain areas such as primary schools and secondary-level humanities, there may well be an over-supply of teachers. Shortages are likely to exist in the areas of science, technology and vocational education, and foreign-language teaching. Indicators on the situation of the supply and demand of teachers in certain subject areas that can be considered are the following:

- the number of unemployed teachers, possibly further specified according to school category and subject area;
- the number of vacancies that have been open for longer than a specific time interval for teachers in the exact sciences and vocational training, relative to the total of teaching positions in the areas concerned;
- the number of teaching periods in a particular subject, taught by unqualified teachers; and
- figures on the stability and mobility of the teaching force.

Basic Financial Conditions

Material and financial conditions of the teaching profession can be captured by means of indicators on:

- gross salary adjusted in several ways to make international comparisons feasible, for instance by using purchasing power parities;
- the above indicators compared with the earnings of comparable employees in other sectors;
- a breakdown of the teaching force at a particular level into those with permanent tenure and those temporarily employed;
- the number of fringe benefits and allowances that teachers are entitled to receive; examples of fringe benefits are: allowances for cost of living, housing/residence, holidays, overtime, remedial teaching or special education; and
- pension and retirement benefits.

Work-load

The work-load of teachers can be captured by looking at official working time, the number of pedagogical obligations teachers have apart from their core job of preparing and carrying out instruction, but also by examining figures on absenteeism of teachers due to illness, and figures on teachers' leaving the profession for reasons other than having reached the age of retirement.

Career Opportunities

As stated before, career opportunities will depend on the situation concerning the supply and demand of teachers in a specific area. Apart from that, career opportunities are likely to be assessed by the teachers concerned in a relative way, that is, by comparing their own position with that of employees in other settings, or of teachers at other levels of the education system. The opportunity structure will further depend on the vertical differentiation of formal responsibilities: in other words, the number of hierarchical positions that teachers can reach through being promoted. More specifically, in this area the following indicators are thought to be relevant:

- the degree of formality of promotion procedures;
- the number of hierarchical levels to which teachers can be promoted (not infrequently this number is just two: teacher and headteacher);
- the magnitude of salary differences between hierarchical levels;
- the proportion of senior posts available in a country;
- the degree of differentiation of salary scales;
- the number of years required to reach the top of the salary scale; and
- facilities to improve professional status through opportunities for further education and training.

The three rationales to guide the selection of teacher indicators discussed in the above have generated ideas for quite a few indicators. Presently, within the framework of the INES project, only a limited set has been selected for further clarification and

operationalisation. Before turning to these, a brief summary of available comparative studies on teachers and national teaching forces will be presented in the next section.

3. Methods Used in the OECD INES Survey on Teachers

Standard Definitions

One of the conditions in comparing data from different countries in a reliable way, is to define the topic the data are about as explicitly and unequivocally as possible. As each country by tradition has its own definitions, it is not always easy to find an overall definition that suits all countries. The definitions used internationally are not always in accordance with definitions used by the central statistical offices in the various countries, and thus not all national statistics are always easy to interpret according to the international definitions.

Procedure

Data were collected through the country representatives in Network C. The data sources on which the figures were based were explicitly mentioned on the answer sheets. For each country, the national co-ordinator of the INES project checked the figures.

4. Discussion of Results

The results of the survey on teachers carried out by Network C of the INES project will be published in the third edition of *Education at a Glance* (OECD, 1995). Here, only the main trends will be discussed, together with some background information that is not included in *Education at a Glance* and an occasional comparison with other international surveys on teachers.

The results of the INES survey on teachers were used to compute indicators on teaching time, duration of teachers' education, teacher compensation and teacher characteristics (age and sex). The data availability for prospective indicators on "stability of the teaching force" and "teacher experience" was so low that they had to be dropped. The number of OECD countries included in the calculation of the indicators varied. Table 3.1 gives an overview of the response with respect to P53, P54, P56 and P58.

Teaching Time

Although an attempt was made to capture both the total official working time of teachers and the teaching time, the concept of official working time appeared to be meaningless or unknown in several countries. The decision had therefore to be made to restrict the scope of the indicator to the time teachers actually teach (in hours per year). Of course this implies a severe reduction with respect to gaining an idea of the total work-load of a teacher. Van Gils (1988) computed the length of the "non-instructional week" as the difference between the total (official) working time of a teacher and the "instructional week". According to his data, there were significant differences in both

Table/Tableau 3.1.

Overview of response in the OECD INES survey on teachers
Vue d'ensemble des réponses dans l'enquête OCDE/INES sur les enseignants

	Teaching time	Teacher education	Teacher compensation	Age and sex
Austria	x	x	x	x
Belgium	x	x	x	x
Finland	x	x	—	x
France	x	x	—	—
Germany	x	x	x	x
Ireland	x	x	x	x
Italy	x	x	x	x
Netherlands	x	x	x	x
New Zealand	x	x	x	x
Norway	x	x	x	—
Portugal	x	x	x	x
Spain	x	x	x	x
Sweden	x	x	x	x
Switzerland	—	—	—	—
Turkey	x	x	x	—
United Kingdom	x	x	x	x
United States	x	x	x	x
Total	16	16	14	13

x = data available.

— = data not available.

the instructional and the non-instructional week among EC countries, the former varying from 17.5 hours (Greece) to 26 hours (the Netherlands), and the latter from 1.75 hours (Denmark) to 14 hours (the Netherlands).

The results of the OECD INES survey likewise showed considerable variance in teaching time between countries, the range being between about 600 and 1 100 hours per year for both ISCED 1 and ISCED 2, while at ISCED level 3 (vocational) this range is even larger: between 612 (Italy, Portugal and Sweden) and 1 692 (Turkey).

In nearly all countries, the number of teaching hours per year in public primary education is higher (mean 858) than in public secondary education (mean 781). The difference between ISCED 2 and ISCED 3 in amount of teaching time is negligible. The mean teaching time at ISCED 3 (general) is 745 hours and at ISCED 3 (vocational) 789.

At a rather abstract level, one could argue that the variance in teaching time among countries indicates that policy-makers have a sizeable margin within which to modify this factor (by increasing or decreasing the school day, the school week or the school year), considering what is apparently possible in other countries. In actual practice, the degree to which teaching time is likely to be adapted by means of policy-measures is of course dependent on the local circumstances. From the point of view of teachers' work-load, one

could take "teaching time" as a proxy indicator of the work-load of a teacher, and tentatively assume that teachers in countries where they have to spend more hours in the classroom have a higher work-load.

Duration of Teachers' Education

Duration of teachers' education was computed as the total number of years of formal education required to become a teacher (thus including the years of primary and secondary education as well as tertiary education specifically providing teacher training). Other aspects of teacher training of which measurement was attempted were the type of arrangement of teacher training and the emphasis on professional (pedagogical) training on the one hand versus subject-oriented training on the other. Vonk (1991) presents results which indicate that among 21 industrialised countries, teacher training is generally arranged concurrently (13 countries). In such a programme, academic (subject-oriented) and professional studies take place in parallel. In three cases (England, Scotland and Northern Ireland), academic and professional studies are integrated, whereas in four other cases teacher training has a consecutive arrangement, professional studies starting after the completion of academic studies (the Federal Republic of Germany – former territory –, France, Ireland and Norway). From Vonk's data it is also evident that in a majority of countries there is an even division of subject-oriented and professional studies within teacher-training programmes. In some countries, France, the Federal Republic of Germany (former territory) and Norway, there is an emphasis on subject-oriented study, whereas for Ireland and Italy the opposite occurs (*i.e.* an emphasis on professional studies).

In the OECD INES survey, data on the type of teacher-training arrangements were incomplete to the degree that it was decided not to publish the results. The data that were obtained on the total duration of teachers' education showed a considerable uniformity across countries. For the pre-primary, primary and secondary levels, teacher education lasts 15 or 16 years in most countries. At primary level the figures range from 13 in Italy to 19 in Germany. The high figure for Germany is partly to be explained by the fact that the time used by new teachers who study teaching through practice has been included.

The relevance to policy-making of these figures on duration of teachers' training may appear to be limited. Duration may not be a key variable, almost certainly not beyond a given threshold. For one thing, the differences do not appear to be dramatic. And, moreover, improving education by means of extra investment in teachers' education takes considerable time and money. Nevertheless, according to various educational and organisational theories (*e.g.* Mintzberg's conception of the professional bureaucracy – Mintzberg, 1979), increasing professionalism through training is the main means of enhancing the functioning of schools. In the long term, it remains to be seen what is more rewarding: thorough teacher education on the one hand, or all types of investment in educational administration (*e.g.* a professional versus a managerial approach) on the other. Hausman and Boyd's (1994) comparison of the educational situation in Germany and the United States is illustrative in this respect.

Teacher Compensation

Table 3.2 shows a comparison between the data collected in the OECD INES survey on primary-school teachers' salaries with figures presented by Van Gils (1988) for EC countries. Minimum and maximum salaries were computed in purchasing power parity rates (PPPs) for EC countries for the year 1985.

From Table 3.2 it can be seen that the purchasing power of primary-school teachers has increased in all five countries for which data are available from both studies. The starting salaries are lowest in the Netherlands in 1985 as well as in 1992; the highest starting salaries are earned in Germany in both years. Maximum salaries have risen relatively strongly in Ireland. The rank order of countries for salaries of starting teachers changed in two out of five cases.

The considerable variation in teachers' salaries among countries indicated in Van Gils' (1988) data was also found among the 16 countries included in the OECD INES survey. For primary education, starting salaries varied from 6 994 PPPs (Turkey) to 23 627 PPPs (Germany), and at secondary level maximum salaries ranged from 12 409 (Turkey) to 42 448 (Austria), measured in PPPs. In order to put these figures into perspective, one would like a referent to which teacher salaries can be compared in each country. Van Gils (1988) compared teacher salaries to average salaries of non-manual industrial workers and manual industrial workers. This comparison indicated that starting primary-school teachers in all the EC countries considered earned less than non-manual industrial workers and that in about half of the countries the same was true in comparison to manual industrial workers.

In the OECD INES survey on teachers, a general welfare index per country was computed by expressing the Gross Domestic Product per head of the population (GDP/h)

Table/Tableau 3.2.
Comparison of salary data in EC countries between 1985 and 1994
*Comparaison des données sur les traitements dans les pays de la CE
entre 1985 et 1994*

	EC 1985		OECD 1992	
	Min.	Max.	Min.	Max.
Belgium	15 274 (4)	25 640 (2)	17 531 (2)	28 582 (2)
Germany	17 979 (5)	266 53 (5)	23 627 (5)	32 464 (4)
Netherlands	13 045 (1)	26 275 (4)	16 819 (1)	30 969 (3)
Ireland	13 600 (3)	25 712 (3)	17 748 (3)	35 985 (5)
Italy	13 285 (2)	18 966 (1)	18 161 (4)	27 852 (1)

The numbers in brackets give the rank orders in a particular year (1 lowest, 5 highest).

Source: Van Gils (1988).

in PPPs. In most countries new teachers at secondary level earned as much as or more than this general welfare index. In Spain and Turkey, the starting salaries of teachers were substantially higher than the GDP/h, which might indicate a relatively high standing of the teaching profession in these countries.

There also appeared to be significant variation with respect to both the difference between minimum and maximum salaries of secondary school teachers and the number of years required to rise from minimum to maximum salary. Assuming equal annual increments on the salary scale, these two findings combined lead to differences in annual increases in salary that range from 170 PPPs in Portugal to 2 523 in the United Kingdom.

A technical issue that may hamper the comparison of gross salaries between countries is non-regular and additional benefits, and the way certain allowances (e.g. family allowances) are made available to teachers in a country. In Tables 3.3, 3.4, 3.5 and 3.6, an overview is given of the amount and nature of non-regular additional benefits.

There is a large variation between countries in the benefits paid in addition to salaries. The percentages vary from 0 per cent of the maximum gross salary in Sweden at all ISCED levels, to 31.4 per cent in New Zealand for teachers in public lower secondary education. The percentage of additional benefits in public lower secondary education is equal to or higher than the percentage for public primary education in all countries.

Additional benefits as a percentage of the maximum gross salary refer to benefits which a full-time teacher with maximum qualifications, at the top of the salary scale, may acquire in addition to the gross maximum salary (of a teacher with maximum qualifications) (year of reference: school year 1991/92).

There are many ways in which policy-makers may use salaries as an instrument to make teaching more or less attractive in a country. The sizeable variation among the OECD countries considered in this survey on most of the relevant aspects of salaries (level, difference between minimum and maximum, number of years required to rise from minimum to maximum) is an indication that, at least theoretically, countries have quite a few options in this respect.

Teacher Characteristics (Age and Sex)

The figures that were obtained in the OECD INES survey concerning the age distribution of teachers show that in almost all countries there is an ageing teaching population. The age group from 40 to 49 is the largest, and there are also sizeable proportions in the 50-60 years age-range. In the near future, when teachers in this age-range retire, shortages are to be expected. A large majority of teachers at pre-primary and primary levels are female. At the second stage of secondary education men and women are about equally represented. There are no indications that this state of affairs is troublesome to countries. In theory, active positive discrimination measures could be thought of to stimulate, for instance, the influx of younger teachers (one possible measure being a reduction of the age at which teachers retire).

Table/Tableau 3.3.

**Additional benefits as a percentage of the maximum gross salary
per year for public primary and lower secondary education**

*Prestations supplémentaires en pourcentage du traitement brut maximum par an
pour l'enseignement public du premier degré et du second degré, 1^{er} cycle*

	ISCED 1 (%)	ISCED 2 (%)	Gross salary highest qualification Top scale	
			11	12
Belgium	0	0	28.582	31.308
Germany	23.3	25.2	32.464	36.119
Netherlands	11.3	-	30.969	-
New Zealand	22.4	31.4	23.412	25.310
Portugal	3.3	5.9	36.155	39.588
Spain	21.7	21.7	30.632	30.632
Sweden	0	0	18.099	19.698
Turkey	6.1	7.2	12.494	12.494

Table/Tableau 3.4.

Specification of additional bonuses
Précisions sur les primes supplémentaires

	Additional bonuses
Belgium	- none
Finland	- if the teacher is participating in school improvement projects and if he or she has special tasks
	- principal and aiding headmaster tasks
Germany	- higher pay group
Netherlands	- none
Portugal	- for special education teachers
	- for principalship (ISCED 2, 3)
	- for teacher trainers
Spain	- for ISCED 3 Heads of Department, General + Vocational, public
	- for ISCED 3 Vocational, private
Sweden	- none
Turkey	- for additional teaching hours
	- for foreign languages upon certification
	- upon the delivery of a child
	- in case of a death in the family
	- for medical expenses
	- as a reward for success
United Kingdom	- degrees/diplomas
	- posts of responsibility
	- principalships
	- for special education teachers
	- Gaeltacht grant, Island allowance
United States	- for sports coaches

Table/Tableau 3.5.

Regular bonuses as part of the salary
Primes régulières faisant partie du traitement

	Regular bonuses
Belgium	- regional bonus - thirteenth month plus social security benefits - holiday bonus
Finland	- holiday bonus (50% of monthly salary)
Germany	- thirteenth month - holiday bonus
Netherlands	- holiday bonus (8%)
Portugal	- thirteenth month - holiday allowance - meals allowance
Spain	- a bonus for each completed three-year period of service
Sweden	- none
Turkey	- for unemployed spouse - family allowance (up to two children) - job difficulties and risks - preparing for the school year (once a year) - for expenditures to the rent (excluding occupants of rent reduced government houses) - as an incentive for savings - as a contribution for house acquisition (excluding occupants of rent reduced government houses)
United Kingdom	- none
United States	- none

Table/Tableau 3.6.

Provision of family allowances
Allocations familiales

Family allowances	BEL	FIN	GER	NET	POR	SPA	SWI	TUR	UK	US	Total
Are provided in a way not connected to salary at all	x			x	x		x				4
Form a regular bonus that is included in the gross salary			x					x	x		3
Are provided in the form of a tax reduction and thus increase net salaries	x	x				x				x	4

5. Conclusion

The four areas for which data on teachers could actually be gathered from the OECD countries surveyed form only a small sub-set of the potentially fruitful data categories that were mentioned in the introductory section of this chapter. In some cases in which data collection was attempted, the response rate was so low that these areas had to be dropped. Data on flows within the stock of teachers, such as the proportion that leave the teaching profession for reasons other than retirement, is a case in point. For the areas on which data *could* be gathered it happened more than once that only part of the information we wished to obtain could be captured. In the case of the work-load of teachers, for instance, we tried to get information on official working hours, time for preparing lessons, marking of papers etc., and teaching hours. Only the latter appeared to be feasible, so it is fair to say that we obtained only partial information on the work-load of teachers in different countries.

Another example is the type of teacher training: either integrated training in the subject matter and in pedagogy, or a consecutive arrangement whereby a period of training in a subject area is followed by a period of general training in "how to teach". Data availability was too limited to use the information for a comparison between OECD countries.

The fact that lack of data limited the scope of the set of teacher indicators that could be constructed had, as a matter of course, also a negative effect on its validity and relevance for policy-making. The fact that for certain areas only partial information was available led to a rather "reductionist" picture, which was the case for both work-load and teacher education.

Notwithstanding these limitations, the evidence that was given on the state of the teaching force in the OECD countries concerned is interesting and may also be expected to have relevance to policy-making. There appear to be substantial differences in the teaching time (contact hours) between countries, and this comparative perspective on the situation in a particular country could stimulate change. The fact that it is uniformity rather than variation that is predominant in the duration of teacher education is also informative. At the same time, the difference between countries that require the lowest number of years of education and the highest number in this area is quite substantial. The data on teacher salaries are probably the most interesting among the set of current indicators. Not only are there substantial differences in the absolute values of minimum and maximum salaries between countries, but the same is true with respect to the difference between minimum and maximum within countries, and the number of years it takes to rise from minimum to maximum salary (the extremes being eight in New Zealand and 45 in Spain). It is also quite interesting to note that in some countries (Turkey and Spain) salaries of new teachers are considerably above the general welfare level. All these factors are relevant with respect to the attractiveness of the teaching profession in a particular country.

Another interesting finding is the fact that, to some extent, the OECD countries concerned have an ageing teaching population (the age group 40-49 accounting for the largest proportion of the teaching force, with a sizeable number aged 50-60). The

outcomes with respect to the sex distribution of teachers at ISCED levels 0 and 1 is less surprising.

In the future, there are two ways in which the data situation affecting indicators could be improved: firstly, by providing for better and perhaps more detailed statistics on characteristics of and flows within the stock of teachers (drop-out, teacher mobility, unemployment, shortages in certain subject areas); and secondly, by collecting school-level data on relevant teacher characteristics, such as special training, more exact estimates of the work-load, the discretion teachers have in carrying out their job, and experience.

References

- BROPHY, J. and GOOD, T.L. (1986), "Teacher behavior and student achievement", in M.C. Wittrock (ed.), *Handbook of Research on Teaching*, Macmillan, New York, pp. 328-357.
- CREEMERS, B.P.M. (1994), "De realisatie van de kwaliteit van het onderwijs", in B.P.M. Creemers (ed.), *Deregulering en de kwaliteit van het onderwijs*, RION, Groningen.
- FERGUSON, R.F. (1991), "Paying for public education: new evidence on how and why money matters", *Harvard Journal on Legislation*, Vol. 28 (2), pp. 465-498.
- FRASER, B.J., WALBERG, H.J., WELCH, W.W., and HATTIE, J.A. (1987), "Syntheses of educational productivity research", *International Journal of Educational Research*, Vol. 43, pp. 163-179.
- HANUSHEK, E.A. (1986), "The economics of schooling: production and efficiency in public schools", *Journal of Economic Literature*, Vol. 24, pp. 1141-1177.
- HAUSMAN, D.C. and BOYD, W.L. (1994), *Riding the Second Wave across the Sea to Europe: School Administration in the Federal Republic of Germany*, Paper presented at the AERA convention, New Orleans.
- HEDGES, L.V., LAINE, R.D., and GREENWALD R. (1994), "Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes", *Educational Researcher*, pp. 5-14, April.
- JONGE, J.F.M. DE, SENNEMA, E.J., and BEYER, W. (1992), *Salarisvergelijking onderwijzend personeel, ambtenaren burgerlijke rijksoverheid*, Research voor Beleid, Leiden.
- LEUNE, J.H.G. (1988), *Beleid inzake taakomvang en taakbelasting van leraren in het voortgezet onderwijs*, IVA, Tilburg.
- Ministerie van Onderwijs en Wetenschappen (1989), *Rijkdom van het onvoltooide: uitdagingen van het Nederlandse onderwijs*, Ministerie van Onderwijs en Wetenschappen, Zoetermeer.
- MINTZBERG, H. (1979), *The Structuring of Organisations*, Prentice Hall, Englewood Cliffs.
- MONK, D.H. (1993), "Subject area preparation of secondary mathematics and science teachers", *Economics of Education Review*, Vol. 13(2), pp. 125-146.
- NEAVE, G. (1992), *The Teaching Nation. Prospects for Teachers in the European Community*, Pergamon Press, Oxford.
- OECD (1990), *The Teacher Today*, Paris.
- SCHEERENS, J. (1990), "School effectiveness research and the development of process indicators of school functioning", *School Effectiveness and School Improvement*, Vol. 1, pp. 61-80.
- SCHEERENS, J. (1992), *Effective Schooling: Research, Theory and Practice*, Cassell, London.

- SMITH, M.S. and O'DAY, J. (1988), *Research into Teaching Quality: Main Findings and Lessons for Appraisal*, OECD working document, Paris.
- VAN GILS, G. (1988), *De arbeidsvoorwaarden van leraren in de EEG* (The conditions of services of teachers in the European Community), Research voor Beleid, Leiden.
- VONK, J.H.C. (1991), "The professional preparation of primary and secondary school teachers in Europe. The analysis of some trends", in J.T. Voorbach, *et al.* (eds.), *Teacher Education 7. Research and Developments on teacher Education in the Netherlands*, Swets and Zeitlinger, Lisse.
- WALBERG, H.J. (1986), "Synthesis of research on teaching", in M.C. Wittrock (ed.), *Handbook of Research on Teaching*, Macmillan, New York.
- WANG, M.C., HAERTEL, G.D., and WALBERG, H.J. (1993), "Toward a knowledge base of school learning", *Review of Educational Research*, Vol. 63, pp. 249-294.

Chapter/Chapitre 4

Curriculum Indicators in International Comparative Research
*Indicateurs relatifs aux programmes scolaires dans la recherche
comparative internationale*

by

Willem J. Pelgrum, Joke Voogt and Tjeerd Plomp
Department of Education, University of Twente, the Netherlands

In recent decades, political and educational discussions have increasingly raised questions about the quality of education. For monitoring the quality of education, programmes for national assessment of educational progress have been implemented in a number of OECD countries. The increasing awareness of the usefulness of these monitoring systems has resulted in a need for indicators which can help in tracing shortcomings and improving educational outcomes. A potentially powerful set of indicators can be found in the curriculum, which can be conceived at three levels. At the macro level of the education system (nation, region), there exists the intended curriculum, at the meso level (school or classroom) the implemented curriculum, and at the micro level (students) the attained curriculum. For each level, specific indicators seem to be relevant. In addition, curriculum antecedents, such as the social background of a student's family or the school neighbourhood, and resources possessed by schools, can be correlated with curriculum contexts to predict curricular outcomes.

In preparing this chapter we reviewed 70 publications on curriculum indicators from research literature. We found 172 different indicators dealing with the curriculum, and for which actual empirical results from national samples are available. From a first overview it appears that there is a lack of consistent terminology in the field of curriculum indicators. We found four global categories of indicators under which the several curriculum indicators can be classified, namely content, time, participants and instructional characteristics. These categories are discussed in depth. The analysis of the validity and reliability of the indicators used shows that in only twenty of the publications reviewed was attention paid to either reliability or validity, or both. We conclude that curriculum indicators have until now not been very well applied in monitoring systems. Steps for further development and research are proposed.

Note de synthèse

Au cours des dernières décennies, les débats sur l'action des pouvoirs publics et l'enseignement se sont de plus en plus fréquemment articulés autour de la qualité de l'enseignement. On a mis en œuvre dans beaucoup de pays des programmes d'évaluation nationale des progrès de l'instruction. Au départ, il s'agissait dans la plupart de ces programmes de mesurer la réussite scolaire d'échantillons nationaux d'élèves dans certaines matières de base. Cependant, à mesure que l'on prenait conscience de l'utilité de ces dispositifs de suivi, on percevait la nécessité de disposer à la fois d'indicateurs capables de repérer les carences, et de moyens de prendre des dispositions afin d'améliorer les résultats de l'enseignement. On peut trouver dans les programmes d'études une série d'indicateurs susceptibles de traiter de la substance même de l'enseignement.

Le programme scolaire peut être envisagé à trois niveaux. Au niveau général (national, régional) du système éducatif, il existe le programme prévu; au niveau moyen (école ou classe), le programme mis en œuvre; et au niveau individuel (élèves), le programme réalisé. A chaque niveau correspondent des indicateurs. Le programme prévu peut être déterminé en examinant les plans des cours, les manuels ou les sujets d'examens. Le programme mis en œuvre peut être mesuré en fonction du contenu des cours, du temps qui leur est imparti ou des stratégies pédagogiques. Le programme réalisé définit les résultats scolaires en mesurant les compétences cognitives et les attitudes des élèves. En outre, on peut établir une corrélation entre ce qui se trouve en amont du programme, c'est-à-dire l'origine sociale de la famille des élèves ou le quartier de l'école, et les ressources dont dispose l'école, d'une part, et de l'autre le contexte des programmes, pour prévoir les résultats.

Dans ce rapport, nous avons étudié 70 publications dont 36 font état des résultats de la recherche à partir de grandes évaluations nationales et internationales et traitent des indicateurs relatifs aux programmes scolaires. Au cours de cette étude, nous avons prêté une attention particulière au programme mis en œuvre, c'est-à-dire que nous avons défini les indicateurs de ce qui est, ou devrait être, enseigné, par et pour qui, et de quelle façon. En outre, on a cherché à en savoir plus au sujet de la fiabilité et de l'exactitude des indicateurs et de l'usage qui pouvait en être fait pour diagnostiquer les carences de l'enseignement.

Au total, nous avons trouvé 172 indicateurs différents qui traitent du programme scolaire et pour lesquels on dispose de véritables résultats empiriques obtenus sur des échantillons nationaux. Il ressort d'un premier examen que l'on manque d'une terminologie cohérente dans le domaine des indicateurs liés au programme scolaire. Par exemple, un indicateur tel que « la possibilité de s'instruire », qui fait état du programme mis en œuvre, se retrouve aussi aux rubriques des « sujets couverts » ou du « chevauchement entre programmes et examens ». Nous avons trouvé quatre catégories globales d'indicateurs dans lesquelles il est possible de classer les divers indicateurs de programme, à savoir le contenu, le temps, les participants et les caractéristiques de l'enseignement dispensé. Ces indicateurs font l'objet d'une description approfondie dans ce chapitre.

L'analyse de la fiabilité et de l'exactitude des indicateurs utilisés montre que l'on ne s'intéresse que dans vingt des publications étudiées à la fiabilité, à l'exactitude, ou aux

deux. La plupart des remarques concernant la fiabilité se réfèrent aux problèmes posés par la collecte des données, tandis que la remarque au sujet de l'exactitude se réfère à l'exactitude prospective, c'est-à-dire au rapport entre l'indicateur et la réussite scolaire. Seuls quelques auteurs analysent en profondeur la fiabilité et l'exactitude.

Il ressort de cette étude que les incidences des indicateurs relatifs aux programmes scolaires sont rares jusqu'à présent et que ces indicateurs ne sont guère appliqués à des fins de diagnostic dans un système de suivi. Ce chapitre propose des dispositions à prendre pour pousser plus loin la recherche et la mise au point dans ce domaine.

*

* *

1. Introduction

In recent decades, political and educational discussions have increasingly raised questions about the quality of education. These discussions arise especially when education budgets are under pressure and a typical question then addressed is what students really learn at school.

The need for information on the quality of education is apparent in many countries throughout the world and has led to many different research initiatives, such as:

- national assessment studies (Beaton, 1987);
- international comparative studies (Husén, 1967);
- national indicator development projects (Stern, 1986); and
- international indicator development projects of the OECD (1992, 1993).

Programmes for national assessments of educational progress have been implemented in many countries (de Landsheere, 1994) and most often labelled with a medical analogy, namely monitoring. Initially, most of these programmes heavily emphasised the measurement of the educational achievement of national samples of students in basic school subjects, such as mathematics, science and the mother tongue. However, the increasing awareness of the usefulness of these monitoring systems has resulted in a need for indicators which can help in tracing the causes of deficits and in finding clues to measures for improving educational outcomes. A potentially powerful set of indicators dealing with the substance of education can be found in the curriculum.

The relationship between curriculum and assessment was recently discussed in the OECD study *Curriculum Reform: Assessment in Question* (OECD, 1993). Nisbet discusses the question to what extent planned assessment can or should become an integral part of monitoring curricula in order to achieve stated objectives or reform itself. He answers this question by categorically stating that once curricula are known and developed, assessment can and should be developed to monitor levels and change. But in addition, Nisbet (in OECD, 1993) rightly states also the reverse, namely that curriculum reform should have a basis in assessment. That is, assessment planned from the beginning as an integral element in monitoring curricula can contribute to the attaining of the

curriculum objectives, including the identification of a compatible teaching and learning style that accomplishes such objectives. This OECD study illustrates the relevance of studying curriculum indicators within the context of monitoring the quality of education through assessment studies, among other measures.

In this chapter we aim to provide an overview of curriculum indicators which have been used in large-scale assessments, and to discuss their strengths and weaknesses.

2. A Theoretical Perspective

Monitoring and improvement of the output of an education system are closely related concepts. One of the most important vehicles for improvement of the output of education is the curriculum. Aoki (1985, p. 1) states that: "In any serious discussion of school improvement, improvement of curriculum is implied. Curriculum improvement, in turn, implies curriculum evaluation." Nisbet (in OECD, 1993) has the same approach, when he speaks of curriculum-based assessment and assessment-led curriculum reform.

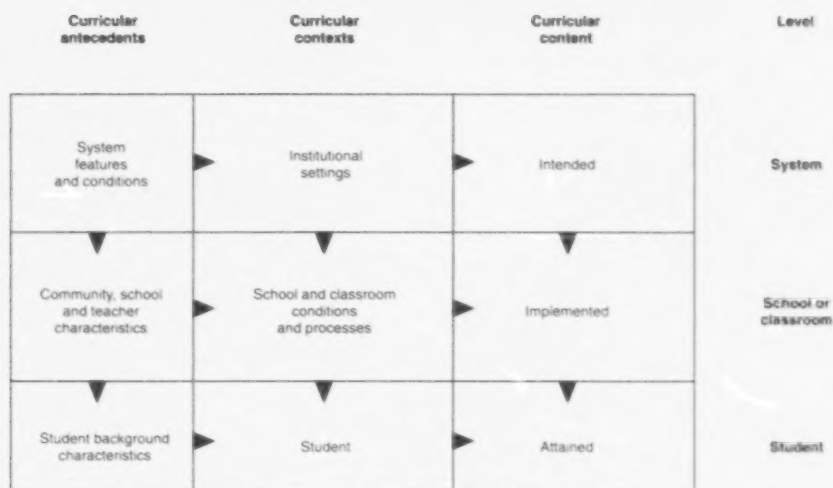
Curricula consist of different components, such as "a) a framework of assumptions about the learner and society; b) aims and objectives; c) content or subject matter with its selection, scope and sequence; d) modes of transactions, for example, methodology and learning environments; and e) evaluation" (Eash, 1985, p. 1147; see also De Corte *et al.*, 1981).

In any discussion of the relationship between curriculum and assessment, the curriculum has to be viewed from three levels (Plomp and Loxley, 1993). What students learn is dependent in part on what teachers choose to teach them. In turn, what teachers choose to teach is dependent on course outlines, official syllabuses, textbooks and the like. These documents are prepared and determined outside the schools on the basis of decisions taken by state (national and/or regional) education authorities, which in turn form the starting point for curriculum developers, textbook authors and publishers in preparing the materials from which schools and teachers make their selections.

Plomp and Loxley (1993), referring to the design of the studies of the International Association for the Evaluation of Educational Achievement (IEA), concluded that curricula can therefore be conceived at three levels: the nation, the school, and the student, where each level can be measured, assessed and compared for degree of overlap and correlation. This is presented in a model in Figure 4.1 (see also Garden, 1987; Travers and Westbury, 1989).

At the macro level of the education system (*i.e.* nation and/or region), there exists the intended curriculum, laid down in official syllabuses, course outlines, or implicit in textbooks, which represent shared conceptions by administrators, teachers, and parents of what is important. At the meso level – the school or classroom – is the intended *and* implemented curriculum, measured by course content, time allocations, instructional strategies, and controlled by teachers. Finally, at the micro level of the student, one finds the attained curriculum, measured by cognitive skills and the attitudes of students which result from teaching and instruction of the prescribed curricula. Although these three aspects of curricula are interlocked some of their features can be separately measured. National curricula can be assessed by examining national outlines, textbooks, examina-

Figure 4.1. Curriculum types applied in IEA studies
Types de programme d'études utilisés dans les études de l'IEA



Note: See Annex.

Source: Plomp and Loxley (1993).

tions, and so on, to determine the range of content offered in any grade. Actual classroom curriculum content can be discerned from teacher and student behaviours based on survey questionnaires which measure what is learned, what is taught, and attitudes towards the teaching-learning process held by students and teachers.

In addition, curriculum antecedents such as the social background of a student's family or the school neighbourhood, as well as resources possessed by schools, can be correlated with curriculum contexts (*e.g.*, institutional setting, instructional processes, and student behaviours) to predict curricular outcomes.

Plomp and Loxley (1993) argue that assessments based on this three-tier approach allow national curricula to be studied at different levels. Besides, when the methodology and instruments are used across many countries, the context can be expanded to the regional or international level. Depending on assumptions of linearity, causality, and the absence of intervening variables, this general model can be used to study a wide range of questions about the quality of education. Possible discrepancies between the three curriculum levels can be studied, for example in the context of educational innovations, by looking at the large number of factors which have been identified as contributing to possible (mis)matches between each pair of curriculum levels: the intended and imple-

mented curriculum; the implemented and attained curriculum; and the intended and attained curriculum.

According to sources such as Fullan (1982, 1991) and Van Velzen *et al.* (1985), the so-called mediating factors can be summarised in the following categories (taken from Van den Akker *et al.*, 1992):

- national (and/or state and/or regional) context;
- characteristics of the school (organisation);
- external support;
- characteristics of the innovation itself.

Van den Akker *et al.* (1992) mention in each category several variables, such as:

National context:

- central legislation and regulations;
- system of policy formation and decision-making in educational affairs;
- time, resources and facilities made available for an innovation;
- proclaimed values and aims of an innovation – the attitudes of politicians and other opinion leaders towards an innovation.

School organisation:

- experiences with earlier innovations;
- role of the school leader(s);
- methods of decision-making;
- available facilities (time, budget, materials);
- internal co-operation and mutual support;
- distribution of information.

External support:

- in-service training;
- assistance with internal staff development;
- coaching and guidance of individual teachers;
- personal contacts with experts and colleagues from other schools.

Innovation characteristics:

- relevance of the innovation for the needs and problems of users;
- clarity of the goals and practical meaning of a proposed change;
- complexity of innovation requirements, especially in teaching behaviour;
- quality and practicality of the innovation products.

Van den Akker *et al.* (1992) qualify the first three categories as conditional for the ultimate use of the products of the innovation by teachers in their daily work. These conditional factors belong to the categories "curricular contexts" and "curricular antecedents" from Figure 4.1.

The conditional factors, together with the innovation characteristics (we will refer to them as mediating factors), can play a role at the different curriculum levels. For

example, in the national context, external support and innovation characteristics may influence the quality of both the intended and implemented curriculum. As all factors directly or indirectly play a role at the level of the implemented curriculum, they influence indirectly the attained curriculum as achieved by students. We therefore assume that the mediating factors may to a large extent determine any discrepancies between the respective levels, or, when applied to monitoring the quality of the curriculum, may contribute to minimising discrepancies between the levels.

From the above one may conclude that the educational research literature contains a large number of factors for potential inclusion in educational monitoring systems. This chapter will not deal with the full range of mediating factors, but will focus on the curriculum indicators outlined in Figure 4.1. Another limitation is that curriculum indicators will not be discussed in the context of bringing about change; the major perspective is monitoring on-going educational processes. A major question to be addressed in the remainder of this chapter is to what extent such indicators have already been used in large-scale assessments and what evidence about their measurement properties is available.

3. The Empirical Scene

A total of 70 documents were selected from the literature possessed by the authors. These were documents pertaining to national and international assessments, as well as documents which according to their titles might be of interest for identifying curriculum indicators. The documents which proved relevant are listed in the references. Although this selection does not pretend to be exhaustive, we believe that the titles listed offer a representative flavour of previous research experiences regarding curriculum indicators.

In analysing the documents, special attention was paid to identifying indicators pertaining to what is or ought to be taught, by and to what participants, and in what way. Moreover, information was sought on the reliability and validity of the indicators and on how the indicator data might be used to diagnose educational shortcomings.

From a first inspection of the results of this review, we learned that most of the more theoretical articles (such as Shavelson *et al.*, 1987; Oakes, 1987; OECD, 1991a; Griffith and Salganik, 1992) do not contain information about the reliability and validity of the proposed indicators and therefore will not be included in the overview below.

In total we found 172 indicators (see Annex to this chapter for a cleaned alphabetical list after the removal of duplicates) which dealt with the curriculum and for which actual empirical results from national samples were available.

One of the first impressions was that there seems to be a lack of consistent terminology in the field of curriculum indicators. For instance, McKnight (1980) refers to "curricular intensity", while others talk about "time spent on teaching and learning". Or, "implemented content" might be found under such labels as "coverage" (Travers and Westbury, 1989), "opportunity to learn" (Keeves, 1992), and "test-curriculum overlap" (de Haan, 1992).

A general observation from the literature review is that there is still a main orientation towards descriptions of results and causal conclusions. Use of curriculum indicators for educational diagnosis still occurs only marginally: very often, conclusions are expressed in terms of countries with low scores on certain indicators, but this does not necessarily imply educational deficits.

The results of our review will be discussed in four global categories, namely content, time, participants, and instructional characteristics.

Content

Forty-three of the indicators (25 per cent) deal with curriculum content. A few of them pertain to the intended level (such as state graduation requirements, number of "new basic" credits recommended, importance of education goals). Most indicators deal with the implemented curriculum. Some of them are global indicators (such as the number of effectively different courses, or average number of subjects in a course), which in themselves do not say much about the actual content. Others, such as the famous IEA "opportunity to learn" measure, are very detailed (such as the percentage of students exposed to subject matter operationalised in a test item). A common characteristic of the content indicators is that they consist of measures indicating the quantity of subject matter covered for several aggregation levels (a whole school subject or (sub)topics within a subject). Different ways of manipulating basic content data may lead to different indicator names. For instance, Keeves (1992) refers to intensity of learning, which is high when all science fields (that is biology, chemistry and physics) are covered. McKnight *et al.* (1987) talk about curricular organisation when measuring at what times various content items are offered. In the same vein, one may use content indicators to study stability and change in the curriculum over time (Hirstein, 1980), variety and diversity (McKnight, 1980), curricular emphasis on task types (Purves, 1992), curricular similarity (Postlethwaite and Wiley, 1992), emphasis on processes (Keeves, 1992), and even curricular validity of tests. The basic question of what measure is most appropriate as an indicator for intended or implemented content was raised by Pelgrum (1990, p. 17):

"One may, in theory, define a particular subject domain by a set of test items. Within this total set of items different sub-sets may be defined. One sub-set would be those items that are considered appropriate operationalisations of the intended curriculum ('intended items'). Another sub-set may consist of those items that refer to the implemented curriculum ('implemented items'), whereas a third set would contain the items that are part of the tests measuring the realised curriculum ('tested items')."

Time

We found 27 (16 per cent) references to indicators of instructional time. Although in itself a time indication does not convey much meaningful information (see also Oakes, 1989), it does indicate the constraints for the amount of content which can be covered and it is a process indicator which in many countries is easily manipulated by school authorities (such as ministries of education and the inspectorate). It has been expressed as the amount of class time in hours (Blank and Dalkilic, 1990; Postlethwaite and Wiley, 1992)

or minutes (Lapointe, 1992; IAEP, 1992) per week, per day (Fey, 1980; Langer, 1990), the total number of hours in a school year (Keeves, 1992), or the hours of instruction the "average" student receives during the school year (OECD, 1991b). Although this measure is sensitive for differences in length of the school year between countries (for which measures exist, such as total number of school days in a country, see Elley, 1992; Purves, 1992), one may wonder whether a relative estimate, such as the percentage of total timetable per course (Travers and Westbury, 1989; Keeves, 1992) does not mask more than it reveals, because the same percentage may result even when the absolute number of hours differs greatly. Some measures, such as time spent on teaching and learning science in a typical class and a typical week (Mullis and Jenkins, 1988), rely on rather sophisticated higher-order mental skills of teacher respondents.

Almost every researcher has probably experienced the difficulties which teachers have with providing time estimates. This does not apply to the amount of time allocated in the timetable, but to how much time is devoted to particular instructional activities.

Participants

Indicators of who is involved in teaching and learning are probably the most pervasive in educational statistics. Most national statistical agencies produce these statistics routinely, and international statistical overviews contain almost without exception statistics in this area. Fifty-eight references (34 per cent) are indicators of this kind, which most often deal with the background characteristics of teachers – such as gender, age and previous schooling – and of students – such as gender and ethnicity. These indicators offer important context information about education systems and are needed for addressing major policy issues, such as equity, minorities and retention. Furthermore, they could be useful from the point of view of educational diagnostics for locating in which segments of society problems are most likely to occur, and for determining whether one country has found better ways of handling certain problems than others. The latter use still hardly occurs in international assessments, although explorations for finding appropriate methods exist (see Postlethwaite and Ross, 1992).

Instructional Characteristics

The implemented curriculum is not only characterised by the previous categories of content and time, but also by instructional characteristics. Indicators of the way instruction is provided cover aspects such as the use of textbooks (Langer *et al.*, 1990) or number of textbooks per student (Elley, 1992; Lundberg and Linnakyla, 1992), use of instructional tools such as calculators and computers (IAEP, 1992; Pelgrum and Plomp, 1993), amount of feedback provided by the teacher (Applebee, 1986), teaching methods, such as lecturing, demonstration, and discussion (Mullis and Jenkins, 1988) or student choice orientation (Keeves, 1992).

4. Reliability and Validity of Curriculum Indicators

Most of the curriculum indicators mentioned above (with the exception of indicators on the characteristics of participants) pretend to describe certain aspects of what is happening in the classroom. They are mainly based on responses from teachers and students to questionnaires, and hence on the assumption that these respondents offer reliable and valid information about the actual situation in the class. The extent to which this assumption holds true determines the meaningfulness of an indicator. Sometimes the operationalisations of curriculum indicators look quite simple to answer (such as "Do you use a textbook?"); in other cases it looks as though respondents are put in the place of the researcher by having to calculate complex estimates (such as the "number of minutes spent on demonstrations in a typical week"). In twenty (29 per cent) of the publications reviewed, comments were made on either the validity or the reliability of the indicators used. In only seventeen (24 per cent), remarks were made about the validity of the indicators used, while in eleven (16 per cent) of the publications, references were made to the reliability of the curriculum indicators. The remarks on validity referred mostly to the predictive validity of the indicator for student achievement (e.g. Keeves, 1992; Scheerens, 1989). The remarks with respect to reliability mostly referred to problems which were encountered during data collection, such as limited data from countries, different ways of drawing samples in the countries involved (e.g. Keeves, 1992) or difficulties with determining the intended curriculum in decentralised education systems (e.g. OECD, 1991b; Rosier and Keeves, 1991). In these twenty publications, 75 different indicators were investigated. However, only five of them were mentioned more than once. The most frequently mentioned curriculum indicator in this respect is "opportunity to learn", which is investigated in sixteen studies (Coverage or Test-Curriculum Overlap).

"Opportunity to learn" is a measure for the implemented curriculum. It is often used in determining the curricular validity of student achievement tests. In eight studies, "opportunity to learn" was measured by an item-based approach.

Measures using an item-based approach to curriculum content appear to be particularly promising, because of their direct focus on the curriculum content of the implemented curriculum and not on indirect measures such as curricular emphasis (Oakes, 1989) or curricular intensity (McKnight *et al.*, 1987) which only refer to time allocated to (parts of) subjects.

There appear to be only a few studies that have examined analyses of the reliability and validity of item-based content measures. These studies will be described briefly below.

Pelgrum (1990) conducted secondary analyses on data sets from the IEA Second International Mathematics Study and expressed the view that:

"The conclusion from the analyses with respect to the validity of implementation ratings, as used in the IEA Second International Mathematics Study and Second International Science Study, is that the results look promising. Teachers gave parallel items almost the same ratings (research question 1). The correlation between the content of the textbook and the ratings of teachers is quite high (research question 2).

The factor structure of the implementation ratings reflects the structure of the Dutch school types (research question 3). The ratings of students and teachers converge reasonably (research question 5). Thus, the validity of the measure is in general quite high. It was, however, also demonstrated that teachers may make serious judgement errors for particular sub-sets of items. For the Dutch data set, it was shown that teachers have difficulties in determining which subject matter was taught in elementary education (research question 4). Although the correlations between student and teacher ratings are generally quite high, comparisons of teacher and student ratings for some countries showed that teachers tend to under-rate the amount of subject matter taught to students. Lack of additional data prevented the investigation of how this phenomenon might be explained.

Furthermore, it was demonstrated that the construct validity of the implementation ratings of the taught/non-taught dimension was not very clear (research question 6). The available data did not exclude the possibility that teachers confound this dimension with other dimensions.

A comparison of comprehensive and non-comprehensive school systems showed that the implementation ratings are more homogeneous in the comprehensive school systems (research question 7) (Pelgrum, 1990, p. 73)."

De Haan (1992) aimed at improving the existing item-based measures for Test-Curriculum Overlap. Most measures for Test-Curriculum Overlap do not take factors into account that may influence it. In De Haan's study, these factors, for example symbols or terminology used, the test format and the difficulty of language, were included in the instrument. Because of the labour-intensive character of this instrument De Haan also developed an alternative approach in which teachers were asked to select those items from an item set which were considered to be appropriate for testing student achievement at a certain time. The latter instrument appears to be particularly useful if it is sufficient to have general information about whether an item is being taught or not. De Haan investigated the reliability, construct and predictive validity of both measures.

From the reliability study it became clear that both instruments were reliable if a dichotomy score (treated/not treated) was used for teachers' judgements. Such judgements appeared to be less convergent if a specific time period had to be indicated (particularly if this time period did not correspond to the practice of the teacher). Construct validity was obtained through establishing the convergence of the instruments for Test-Curriculum Overlap with other measures such as textbook analysis and student-based analysis of Test-Curriculum Overlap. The correlations between the different approaches for groups of items or groups of teachers appear to be quite high, which indicates that the different measures for Test-Curriculum Overlap can be compared. Additionally, it was established whether both instruments for Test-Curriculum Overlap detect changes in the implemented curriculum over time. This could be established particularly with the comprehensive instrument. With the comprehensive instrument it was also possible to show whether items were deviating on one or more of the possible factors influencing Test-Curriculum Overlap. For determining the predictive validity of both measures an analysis was made of whether student achievement differed between taught and non-taught items. One of the major conclusions from this study is that it is

doubtful whether curriculum indicators referring to opportunity to learn (in terms of whether an item is taught) are good predictors of student achievement.

The use of item-based measures for determining curriculum content was also applied in a study conducted by Voogt and Vlas (1994). The study was aimed at determining deficiencies in students' mathematics skills when they entered higher education. The core of the study was based on the development of a set of mathematical items which describe both the intended, the implemented and the attained curriculum for mathematics at the conclusion of secondary education, and the attained curriculum desirable for entering tertiary education. The study started with a content analysis of mathematics in secondary education and an analysis of the desired mathematics content in certain sectors in tertiary education. This resulted in a description of mathematics content in several sub-domains.

During all stages of the description of the curriculum domains and the development of an item set based on these domains, mathematics experts from secondary and tertiary education were involved and judged the content validity of the item set.

In the study, the whole item set ($N = 101$) was administered to a small group of experts on mathematics in secondary education, who had to rate whether the item represented the intended content of the secondary education mathematics curriculum. A group of about 100 mathematics lecturers from three sectors of higher education (technical, economics, and teacher training) rated whether the item represented the desired attained curriculum: the curriculum content necessary for coping successfully with mathematics in the first year of higher education.

A sub-set of the item set ($N = 37$) was administered at the start of the course to a sample of first-year students in higher education from the three sectors (technical, economics, teacher training) involved in the study. Students not only had to take the test but also to indicate whether the item was taught in secondary education. In this way the attained and implemented curriculum was measured. The reliability of the instruments was determined with the α -coefficient of internal consistency. It appeared to be satisfactory (students' test 0.91, student-based OTL 0.89, higher education teachers' judgements 0.97). Moreover, the approach appeared to be successful in determining discrepancies between the various perspectives on the curriculum.

5. Summary, Conclusions and Recommendations

Quite a number of different curriculum indicators have been used in large-scale international educational assessments. However, the conclusions or implications drawn from these indicators have usually been quite modest, usually descriptive and looking more like window-dressing than a serious attempt at integrating a diagnostic tool into a monitoring system. From many publications, the impression arises that certain indicators are seen more as a goal in themselves than as part of a complex and consistent monitoring system.

This points to a lack of conceptual sophistication in the monitoring function of international assessments, which is not strange given the fact that monitoring of social and educational aspects of a society is a fairly new activity in the development of mankind. In this context it is understandable that there is a lack of information about the

validity of many of the indicators described in the literature, because validity can only be determined when there is a clear definition of the purposes for which an indicator is to be used.

In order to contribute to the improvement of the existing situation, we shall now sketch some steps for a future agenda for further development and research in this area.

Following a medical analogy, one might describe educational monitoring as a facility which constantly registers the critical functions of an education system. The most important question is how students develop in terms of their knowledge, skills, and attitudes. Measures in this area constitute the heartbeat of the education system, and should at the same time make up the heart of the monitor. This is the most crucial, but also the most difficult part to construct, because it needs to contain a model of "good functioning" of the education system which performs a signalling function, generating warnings when outcomes differ from what was expected. This expectancy needs to rely on comparisons, for instance with the outcomes monitored in the past or in relevant reference groups (for example, other countries). It requires the development of a methodology for determining whether students, schools, districts or countries perform less well than may be expected on the basis of what constitutes their intended curriculum and their educational environment. Although some first exploratory steps in developing such a methodology at the international level can be found in Postlethwaite and Ross (1992), we are still far away from an accepted method which integrates the feedback mechanisms at the different educational levels.

Further development should involve co-operation between monitoring activities at school, national, and international levels. This would require international co-operation over a long period of time and should probably first be set up on a small scale as a feasibility study. It requires a flexible supranational co-ordination mechanism which tries to establish standardisation of approaches by, for instance, developing a handbook of indicators and technical standards for the collection of high-quality data, establishing a funding structure for co-funding studies that use these indicators in in-depth studies of reliability and validity issues, developing specifications of reporting formats and sampling methods, and stimulating the inclusion of results in databases integrated into Internet.

If we qualify the first aspect of an educational monitor as the signalling function, the second aspect could be labelled the diagnostic function. This function should generate a description of the location and the possible causes of malfunctioning, so that educational actors may find clues to measures which might improve the outcomes. For this purpose the monitor should contain reliable and valid process indicators, which have been shown to correlate strongly with educational effectiveness and change. It would be very important to include indicators reflecting intervention programmes which are undertaken for improving educational outcomes. Such measures would reflect the extent to which the interventions have actually been implemented and could be used to analyse (longitudinally) how successful these interventions were.

The question may arise, why at the international level monitors with characteristics such as those sketched above still do not exist. Besides the complexity of creating such a system, probably one of the prominent reasons is that most international assessment

activities have arisen in an academic context, and were never designed for monitoring. This has resulted, among other things, in a lack of longitudinal components and an overlap between studies. Furthermore, the costs and benefits of such monitoring are not yet clear. The ultimate challenge is to demonstrate scientifically that international monitoring has added value in enhancing educational productivity.

References

- AKKER, J.J.H. Van den, KEURSTEN, P., and PLOMP, T.J. (1992), "The integration of computer use in education", *International Journal of Educational Research*, Vol. 17 (1), pp. 65-76.
- AOKI, T. (1985), "Interests, knowledge and evaluation: alternative approaches to curriculum evaluation", *Journal of Curriculum Theorizing*, Vol. 6 (4), pp. 27-44.
- APPLEBEE, A. (1986), *The Writing Report Card: Writing Achievement in American Schools*, Educational Testing Service, Princeton, NJ.
- BEATON, A.E. et al. (1987) *Implementing the New Design: the NAEP 1983/84 Technical Report*, National Assessment of Educational Progress, Princeton, NJ.
- BLANK, R.K. (1989), *Development of a 50-State System of Educational Indicators: Issues of Design, Implementation, and Use*, State Education Assessment Center, Washington, D.C.
- BLANK, R.K. and DALKILIC, M. (1990), *State Indicators of Science and Mathematics Education*, State Education Assessment Center, Washington, D.C.
- CASSIDY, W. and BOGNAR, C.J. (1990), *More than a Good Idea: Moving from Words to Action in Social Studies*, Ministry of Education, Victoria, British Columbia.
- CORTE, E. de, GEERLIGS, C.T., LAGERWEIJ, N.A.J., PETERS, J.J. and VANDENBERGHE, R. (1981), *Beknopte didaxologie (Concise didactics)*, Wolters Noordhoff, Groningen.
- EASH, M.J. (1985), "Curriculum components", in Husén, T. and Postlethwaite, T.N. (eds), *The International Encyclopedia of Education*, Pergamon Press, Oxford.
- ELLEY, W.B. (1992), *How in the World do Students Read?*, IEA Secretariat, The Hague.
- FETLER, M.E. (1989), *Assessing Educational Performance: California's School Quality Indicator System*, PACE, University of California at Berkeley, Berkeley, CA.
- FEY, J.T. (1980), "The United States' NSF studies of mathematics education", in *Comparative Studies of Mathematics Curricula - Change and Stability 1960-1980*, Institut für Didaktik der Mathematik, Bielefeld.
- FULLAN, M. (1982), *The Meaning of Educational Change*, Teachers College, New York.
- FULLAN, M.G., with STIEGELBAUER, S. (1991), *The New Meaning of Educational Change*, Teachers College, New York.
- GARDEN, R.A. (1987), *The IEA Study of Reading Literacy*, IEA, Hamburg.
- GRIFFITH, J.E. and SALGANIK, L. (1992), "The OECD international education indicators project: Network C, Features of schools and school systems", Paper presented at the annual meeting of the American Education Research Association, San Francisco, CA.

- HAAN, D.M. de (1992), *Measuring Test-Curriculum Overlap*, Doctoral thesis, University of Twente, Enschede.
- HIRSTEIN, J. (1980), "From Rouaymont to Bielefeld: a twenty year cross-national survey of the content of school mathematics", in *Comparative Studies of Mathematics Curricula - Change and Stability 1960-1980*, Institut für Didaktik der Mathematik, Bielefeld.
- HUSÉN, T. (ed.) (1967), *International Study of Achievement in Mathematics: A Comparison of Twelve Countries*, Volumes I and II, Almqvist and Wiksell, Stockholm/John Wiley, New York.
- IAEP (1992), *Learning Mathematics*, Educational Testing Service, Princeton, NJ.
- KEEVES, J.P. (1992a), *Learning Science in a Changing World: Cross-National Studies of Science Achievement 1970-1984*, IEA Secretariat, The Hague.
- KEEVES, J.P. (1992b), *The IEA Study of Science III: Changes in Science Education and Achievement: 1970 to 1984*, Pergamon Press, Oxford.
- LANDSHEERE, G. de (1994), *Le pilotage des systèmes d'enseignement*, Edition de Boeck, Brussels.
- LANGER, J.A. et al. (1990), *Learning to Read in our Nation's Schools: Instruction and Achievement in 1988 at Grades 4, 8 and 12*, Educational Testing Service, Princeton, NJ.
- LAPORTE, A.E. et al. (1989), *A World of Differences*, Educational Testing Service, Princeton, NJ.
- LAPORTE, A.E. et al. (1992), *Learning Science*, Educational Testing Service, Princeton, NJ.
- LUNDBERG, I. and LINNAKYLA, P. (1992), *Teaching Reading around the World*, IEA Secretariat, The Hague.
- MCKNIGHT, W. and CURTIS, C. (1980), "Classroom processes in mathematics instruction: on characterizing the implemented curriculum", in *Comparative Studies of Mathematics Curricula - Change and Stability 1960-1980*, Institut für Didaktik der Mathematik, Bielefeld.
- MCKNIGHT, W. and CURTIS, C. (eds.) (1987), *The Underachieving Curriculum: Assessing US School Mathematics from an International Perspective*, Stipes Publishing Company, Illinois.
- MEDRICH, E.A., and GRIFFITH, J.E. (eds.) (1992), *International Mathematics and Science Assessments: What Have We Learned?*, US Department of Education, Washington DC.
- Ministry of Education (ed.) (1991), *British Columbia Assessment of Science. Technical Report I: Classical Component*, Victoria, British Columbia.
- MULLIS, I. and JENKINS, L. (1988), *The Science Report Card: Elements of Risk and Recovery*, Educational Testing Service, Princeton, NJ.
- OAKES, J. (1987), *Conceptual and Measurement Issues in the Construction of School Quality Indicators*, RAND Corporation, Santa Monica.
- OAKES, J. (1989), "What educational indicators? The case for assessing the school context", *Educational Evaluation and Policy Analysis*, Vol. 11 (2), pp. 181-199.
- OECD (1991a), *The Task of Identifying International Educational Indicator*, INES Working Paper, Paris.
- OECD (1991b), *Draft Technical Handbook for Education Indicators*, General Assembly on International Educational Indicators, Paris.
- OECD (1992, 1993), *Education at a Glance: OECD Indicators* (bilingual), CERI, Paris.
- OECD (1993), *Curriculum Reform: Assessment in Question*, Paris.

- PELGRUM, W.J., (1990), *Educational Assessment: Monitoring, Evaluation and the Curriculum*, Academisch Boeken Centrum, De Lier.
- PELGRUM, W.J. and PLOMP, T.J. (eds). (1993), *The IEA Study of Computers in Education: Implementation of an Innovation in 21 Education Systems*, Pergamon Press, Oxford.
- PLOMP, T.J. and LOXLEY, W. (1993), *The Contribution of International Comparative Research to Curriculum Reform*, Paper presented at OECD Conference on "The Curriculum Redefined", Paris, 5-8 April.
- POSTLETHWAITE, T.N. and ROSS, K.N. (1992), *Effective Schools in Reading: Implications for Educational Planners: an Exploratory Study*, IEA Secretariat, The Hague.
- POSTLETHWAITE, T.N. and WILEY, D.E. (1992), *The IEA Study of Science II: Science Achievement in Twenty-Three Countries*, Pergamon Press, Oxford.
- PURVES, A.C. (ed.) (1992), *The IEA Study of Written Composition II: Education and Performance in Fourteen Countries*, Pergamon Press, Oxford.
- ROBITAILLE, D.F. (ed.) (1990), *The 1990 British Columbia Mathematics Assessment. Technical Report 1990*, Ministry of Education, Victoria, British Columbia.
- ROSIER, M. (1980), "The second IEA mathematics study in Australia", in *Comparative Studies of Mathematics Curricula - Change and Stability 1960-1980*, Institut für Didaktik der Mathematik, Bielefeld.
- ROSIER, M.J. and KEEVES, J.P. (eds.) (1991), *The IEA Study of Science: Science Education and Curricula in Twenty-Three Countries*, Pergamon Press, Oxford.
- SCHEERENS, J. (1989), *Process-indicators of School-Functioning: a Selection Based on the Research Literature on School Effectiveness*, INES Working Paper, OECD, Paris.
- SHAVELSON, R., MCDONNELL, L., OAKES, J. and CAREY, M. (1987), *Indicator Systems for Monitoring Mathematics and Science Education*, RAND Corporation, Santa Monica.
- STERN, J.D. (1986), *The Education Indicators Project in the U.S. Department of Education*, US Department of Education, Washington DC.
- STERN, J.D. and HALL, P.R. (1987), *Education Indicator Development at the Federal Level*, US Department of Education, Washington DC.
- TRAVERS, K.J. (1980), "The second international mathematics study: an overview", in *Comparative Studies of Mathematics Curricula - Change and Stability 1960-1980*, Institut für Didaktik der Mathematik, Bielefeld.
- TRAVERS, K.J. and WESTBURY, I. (1989), *The IEA Study of Mathematics I: Analysis of Mathematics Curricula*, Pergamon Press, Oxford.
- VELZEN, W.G. van et al. (1985), *Making School Improvement Work: a Conceptual Guide to Practice*, ACCO, Leuven.
- VOOGT, J.M. and VLAS, H.W.J. (1994), *Aansluiting HAVO-HBO: gerealiseerde en gewenste wiskunde kwalificaties* (Link-up general secondary education-higher vocational education: realized and desired mathematics qualifications), OCTO, Twente University, Enschede.
- WESTBURY, I. (1980), "Change and stability in the curriculum: an overview of the questions", in *Comparative Studies of Mathematics Curricula - Change and Stability 1960-1980*, Institut für Didaktik der Mathematik, Bielefeld.

Annex

Overview of indicators and sources

Content indicators

- Access to books in school – (Purves, 1992)
- Amount of/Number of tasks – (Purves, 1992)
- Availability of topics – (Hirstein, 1980)
- Average number of subjects in course – (Postlethwaite and Wiley, 1992)
- Breadth of content – (Medrich and Griffith, 1992)
- Coverage – (Travers and Westbury, 1989)
- Curricular control – (Postlethwaite and Wiley, 1992)
- Curricular emphasis on specific subtopics – (Purves, 1992)
- Curricular emphasis on task types – (Purves, 1992)
- Curricular emphasis – (Oakes, 1989)
- Curricular goals – (Purves, 1992)
- Curricular intensity – (McKnight, 1987; Travers, 1980)
- Curricular organisation – (McKnight, 1987)
- Curricular similarity – (Postlethwaite and Wiley, 1992)
- Curriculum validity of tests – (Keeves, 1992; Rosier and Keeves, 1991)
- Depth of content – (Medrich and Griffith, 1992)
- Differentiation in curriculum – (Medrich and Griffith, 1992)
- Differentiation – (McKnight, 1980)
- Emphasis on (cognitive) skills – (Cassidy and Bognar, 1990)
- Emphasis on content – (Rosier and Keeves, 1991)
- Emphasis on goals – (Ministry of Education, 1991)
- Emphasis on process objectives – (Rosier and Keeves, 1991)
- Emphasis on processes – (IAEP, 1992; Keeves, 1992)
- Emphasis on topics – (Cassidy and Bognar, 1990); Keeves, 1992)
- Extent of coverage/degree of emphasis on content – (Keeves, 1992)
- Importance of education goals – (Stern and Hall, 1987)
- Intensity of learning – (Keeves, 1992)
- Level of difficulty – (Medrich and Griffith, 1992)
- Linking of goals and content – (McKnight, 1987)
- Locus of decision-making – (OECD/INES, 1991)
- New Basics curriculum – (Stern and Hall, 1987)
- Number of effectively-different courses – (Travers and Westbury, 1989)
- Opportunity to learn – (Keeves, 1992; Lapointe, 1989; McKnight, 1987; Medrich and Griffith, 1992; Ministry of Education, 1991; Mullis and Jenkins, 1988; Pelgrum, 1990; Postlethwaite

and Wiley, 1992; Purves, 1992; Robitaille, 1990; Rosier, 1980; Travers and Westbury, 1989; Scheerens, 1989; Travers, 1980)

Percentage of schools that emphasise a particular subtopic – (IAEP, 1992)

Place and variability of topics in curriculum – (Hirstein, 1980)

Quality of school curricula/content covered – (Scheerens, 1989)

Stability and change – (Hirstein, 1980)

Standards of good achievement – (Purves, 1992)

State graduation requirements by percent of students in grade 9-12 – (Blank and Dalkilic, 1990)

Test-curriculum overlap – (Haan, 1992)

Timing of instruction in particular topics and subjects – (Medrich and Griffith, 1992)

Topic coverage – (OECD/INES, 1991)

Variety and diversity – (McKnight, 1980)

Time indicators

Amount of instructional preparation time – (Medrich and Griffith, 1992)

Average length of class period per subject – (Purves, 1992)

Class time spent on instruction – (Langer, 1990)

Curriculum time allotment – (Robitaille, 1990)

Curriculum time – (OECD/INES, 1989)

Effective learning/time on task – (Scheerens, 1989)

Hours of instruction per course – (Postlethwaite and Wiley, 1992)

Hours of instruction – (OECD/INES, 1991)

Instruction time – (Elley, 1992; Lapointe, 1992; Lundberg and Linnakyla, 1992; Oakes, 1989)

Length of compulsory education – (Purves, 1992)

Length of school year – (Purves, 1992)

Number of hours of instruction per course – (Travers and Westbury, 1989)

Number of school days – (Elley, 1992)

Percentage of total time given to course – (Keeves, 1992)

Percentage of time spent at different cognitive levels – (Ministry of Education, 1991)

Percentage of total timetable per course – (Travers and Westbury, 1989)

Teacher/hour per student – (OECD/INES, 1989)

Teaching time for reading – (Lundberg and Linnakyla, 1992)

Time for homework – (Lundberg and Linnakyla, 1992)

Time learning science – (Keeves, 1992)

Time on course – (Blank and Dalkilic, 1990; Fey, 1980; IAEP, 1992; Medrich and Griffith, 1992; Robitaille, 1990)

Time on task – (OECD/INES, 1991)

Time spent on learning – (Keeves, 1992)

Time spent on teaching and learning – (Mullis and Jenkins, 1988)

Time spent on teaching science – (Ministry of Education, 1991)

Time spent on teaching – (Elley, 1992)

Time spent on the study of the course – (Keeves, 1992)

Instructional characteristics indicators

Assessment of lower-order skills – (Lundberg and Linnakyla, 1992)

Availability of books – (Elley, 1992)

Availability, use and adequacy of resources and facilities – (Ministry of Education, 1991)
 Change in approach to evaluation – (Robitaille, 1990)
 Change in approach to teaching – (Robitaille, 1990)
 Change in teaching practices – (Langer, 1990)
 Classroom activities – (Lapointe, 1989)
 Classroom practices – (Robitaille, 1990)
 Comprehension instruction – (Lundberg and Linnakyla, 1992)
 Degree of achievement-oriented policy – (Scheerens, 1989)
 Degree of evaluation and monitoring of pupils/progress – (Scheerens, 1989)
 Degree of integration – (Ministry of Education, 1991)
 Differentiated and integrated learning – (OECD/INES, 1991)
 Encouragement to read – (Lundberg and Linnakyla, 1992)
 Emphasis on assessment – (Lundberg and Linnakyla, 1992)
 Emphasis on evaluation methods – (Cassidy and Bognar, 1990; Ministry of Education, 1991)
 Emphasis on phonics – (Lundberg and Linnakyla, 1992)
 Emphasis on teaching strategies – (Cassidy and Bognar, 1990)
 Frequency of instructional techniques – (Fey, 1980)
 Grouping practices – (Oakes, 1989)
 Instructional format as a percentage of class time – (Ministry of Education, 1991)
 Instructional materials – (Langer, 1990)
 Instructional strategies – (Ministry of Education, 1991)
 Intensity of teaching – (Keeves, 1992)
 Materials used in instruction – (Purves, 1992)
 Methods of teaching – (Keeves, 1992)
 Number of textbooks – (Elley, 1992; Lundberg and Linnakyla, 1992)
 Orientation on student interests – (Lundberg and Linnakyla, 1992)
 Percentage of students who do exercises/experiments themselves – (IAEP, 1992)
 Percentage of students who listen to course – (IAEP, 1992)
 Percentage of students who solve problems in groups – (IAEP, 1992)
 Reading aloud to students – (Lundberg and Linnakyla, 1992)
 Reinforcement – (Scheerens, 1989)
 Student-oriented reading – (Lundberg and Linnakyla, 1992)
 Structured/direct learning – (Scheerens, 1989)
 Student activities as a percentage of class time – (Ministry of Education, 1991)
 Teacher feedback – (Applebee, 1986)
 Teacher-centred instruction – (Lundberg and Linnakyla, 1992)
 Teaching carried out – (OECD/INES, 1989)
 Teaching emphasis – (Keeves, 1992)
 Teaching materials – (IAEP, 1992)
 Teaching practices – (Mullis and Jenkins, 1988)
 Teaching techniques – (Travers, 1980)
 Use of equipped classrooms – (Postlethwaite and Wiley, 1992)

Participants indicators

Average annual compound growth rate of enrolment – (Westbury, 1980)
 Average size of schools – (Purves, 1992)
 Characteristics of teachers – (Keeves, 1992)

Class size - (Blank and Dalkilic, 1990; Elley, 1992; Lundberg and Linnakyla, 1992; McKnight, 1987; Medrich and Griffith, 1992; Postlethwaite and Wiley, 1992)
 Course enrolment - (Blank, 1989; Fey, 1980)
 End of secondary education - (Purves, 1992)
 Enrolment in secondary education - (Westbury, 1980)
 Enrolments in high-school courses - (Blank and Dalkilic, 1990)
 Freshmen enrolled in remedial courses - (Stern, 1986)
 Gender of teacher - (Blank and Dalkilic, 1990)
 Learning-teaching ratio - (OECD/INES, 1991)
 Minority teachers by minority students - (Blank and Dalkilic, 1990)
 Number of post-secondary courses taken - (Robitaille, 1990)
 Number of schools by number of teachers - (Blank and Dalkilic, 1990)
 Number of science content courses taken - (Ministry of Education, 1991)
 Number of science methods courses taken - (Ministry of Education, 1991)
 Number of students in private schools - (Purves, 1992)
 Number of teachers - (Westbury, 1980)
 Number of years of education for primary school teachers - (Purves, 1992)
 Participation rate - (Keeves, 1992)
 Percentage enrolled as first-time freshmen at University - (Fetler, 1989)
 Percentage graduates completing the a-f course requirements - (Fetler, 1989)
 Percentage of schools where classes are based on ability - (IAEP, 1992)
 Percentage of schools where course teacher has post-sec. educ. - (IAEP, 1992)
 Percentage of schools where teacher teaches just course sub - (IAEP, 1992)
 Percentage of students in each grade taking (advanced)course - (Blank and Dalkilic, 1990)
 Percentage of students who take a course test - (IAEP, 1992)
 Percentage of teachers assigned out-of-field - (Blank and Dalkilic, 1990)
 Percentage of teachers that attended an in-service programme - (Cassidy and Bogner, 1990)
 Percentage of teachers under 30 and over 50 years of age - (Blank and Dalkilic, 1990)
 Percentage of teachers with college major in field - (Blank and Dalkilic, 1990)
 Percentage of an age group studying - (Postlethwaite and Wiley, 1992)
 Percentage of schools teaching course - (Keeves, 1992)
 Percentage of students in programme - (Langer, 1990)
 Pre-school enrolment - (Elley, 1992)
 Primary assignment of teachers of Mathematics and Science - (Blank and Dalkilic, 1990)
 Pupil-teacher ratio - (Postlethwaite and Wiley, 1992; OECD/INES, 1989; OECD/INES, 1991; Purves, 1992; Scheerens, 1989)
 Ratio of those supporting and those supported by schools - (Westbury, 1980)
 Remedial teaching or developmental courses - (Stern and Hall, 1987)
 Retention rate - (Keeves, 1992)
 Retentivity - (Travers and Westbury, 1989)
 School enrolment - (Medrich and Griffith, 1992)
 School entry age - (Purves, 1992)
 Selectivity - (Purves, 1992)
 Sex of teacher - (Elley, 1992)
 Starting age instruction - (Elley, 1992)
 Status of teacher - (McKnight, 1987)
 Teacher experience - (OECD/INES, 1989; Scheerens, 1989)
 Teacher qualification - (OECD/INES, 1989; Scheerens, 1989)
 Teachers assigned 50 per cent or more course - (Blank, 1989)
 Teachers by certification status - (Blank, 1989)

Teachers' certification - (Mullis and Jenkins, 1988)
Teaching course experience - (Ministry of Education, 1991)
Teaching experience - (Robitaille, 1990; Lundberg and Linnakyla, 1992)
Time at which learning course begins - (Keeves, 1992)
Years of post-secondary education - (Cassidy and Bognar, 1990)
Years of required post-secondary training - (Purves, 1992)
Years of teaching experience - (Cassidy and Bognar, 1990; Mullis and Jenkins, 1988)

**Process Indicators on School Functioning and the Generalisability
of School Factor Models Across Countries**

***Indicateurs de processus sur le fonctionnement de l'école
et la possibilité d'application transnationale des modèles de facteurs
scolaires***

by

Bert Creemers

Department of Education, University of Groningen, the Netherlands

Education indicators about the state of the whole education system, or important components of it, at first addressed input and outcomes of education. Currently, the development of indicators at the process level has been given increased attention. Research on processes at the classroom and school levels can reveal factors that contribute to the outcomes of education. These factors can be transformed into process indicators on school functioning or effectiveness. At the national or district level, decision-makers are more interested in broad process indicators that provide information about general factors that can explain differences in outcomes. At the local level, and even more at the school level, there is interest in more refined information about relevant factors at the school and classroom levels. A specific problem connected with process indicators is the generalisability of the factors. In national studies, differences have been found in the results with respect to process factors. This is even more the case in international studies.

Within the International School Effectiveness Research Programme (ISERP), a conceptual model has been developed, and empirical data are being collected. ISERP also looks at the variation between countries, in order to obtain information on the generalisability of process factors between countries. The preliminary results of the first ISERP study are described in this chapter.

*

* *

Note de synthèse

Les indicateurs de l'enseignement qui fournissent des renseignements au sujet de l'état de tout ou partie du système d'enseignement traitent à l'origine des résultats de l'enseignement et des moyens mis en œuvre pour les obtenir. A l'heure actuelle, on prête plus d'attention à l'élaboration d'indicateurs de processus. A cet égard, la recherche sur l'efficacité de l'enseignement qui tient compte des facteurs susceptibles d'expliquer les différences des résultats scolaires peut fournir une importante information. La recherche sur les processus au niveau de l'école et de la classe peut mettre en évidence certains facteurs qui contribuent aux résultats de l'enseignement. Ces facteurs peuvent à leur tour être transformés en indicateurs de processus concernant le fonctionnement ou l'efficacité de l'école.

Les divers niveaux du système d'enseignement ne s'intéressent sans doute pas autant les uns que les autres à ces indicateurs de processus. Au niveau du pays ou de la circonscription scolaire, les décideurs s'intéressent davantage à des indicateurs de processus généralement définis qui renseignent au sujet de certains facteurs qui se situent au niveau de la classe ou de l'école et qui peuvent expliquer les différences de résultats. Au niveau local, et plus encore au niveau de l'école, on s'intéresse à une information plus fine concernant les facteurs qui relèvent du niveau de l'école et de la classe.

L'élaboration d'une théorie de l'efficacité de l'enseignement et la recherche connexe mettent en lumière l'importance d'un grand nombre de facteurs de processus et de variables qui, séparément, expliquent une part minime de la variance. Certaines de ces variables peuvent être exploitées ensemble pour définir de façon empirique des indicateurs de processus plus généraux, tels que la possibilité de s'instruire, la qualité de l'enseignement et le temps consacré à la tâche.

L'un des problèmes propres aux indicateurs de processus est la possibilité de généralisation des facteurs. Dans les études nationales, on trouve des différences entre les résultats relatifs aux facteurs de processus. Ces différences sont encore plus sensibles dans les études internationales. Elles tiennent peut-être à la conception et à la mise en œuvre des études internationales (et nationales). En outre, des différences entre les contextes propres aux pays peuvent sans doute expliquer les différences entre les contributions des facteurs de processus d'un pays à l'autre.

Dans le cadre du Programme international de recherche sur l'efficacité de l'école (International School Effectiveness Research Programme ou ISERP), on a mis au point un modèle théorique qui contient d'importants facteurs relatifs aux différents niveaux de l'enseignement et on a rassemblé, au titre des projets de recherche, des données qui servent à la vérification empirique de la théorie. Ensuite, l'ISERP étudie les variations entre pays afin d'en savoir plus au sujet des possibilités de généralisation transnationale des facteurs de processus. Les résultats provisoires de la première étude de l'ISERP sont décrits dans ce chapitre. Ainsi, l'ISERP, comme d'autres projets internationaux, peut contribuer à la mise au point des indicateurs de processus et offrir des renseignements sur la façon dont doivent être mesurés au niveau des variables ces indicateurs nécessaires qui font l'objet d'une définition générale. La qualité de l'enseignement, la possibilité de s'instruire et le temps consacré à l'étude semblent contenir un grand nombre de

variables qui doivent être mesurées d'une façon particulière. Le temps consacré à l'étude paraît être un indicateur relativement facile; en revanche, la possibilité de s'instruire et, plus encore, la qualité de l'enseignement, seront très difficiles à transformer en indicateurs fiables et comparables sur le plan international.



1. Introduction: The Value of Process Indicators

An education indicator says something about the functioning of the education system and provides information about the state of affairs of the whole education system or an important component of it. Because education indicators address the system as a whole, they do not provide detailed or precise information. The objective of the education indicator is to tell something about the quality of the system and to enable measures to be taken for improvement. It is therefore to collect data that can be compared to a point of reference, the standard for a specific indicator, in such a way that an evaluation and a judgement of quality are possible.

The system of education indicators does not have to be exhaustive. It can include every possible indicator, but for the purpose of evaluation and quality judgement it is necessary that there is a relationship between the indicator and the ultimate criterion of education: student achievement (output indicator). Policy-makers are interested in education indicators because in this way they are provided with concise and relevant information about the system. This implies that they are not interested in fine, precise process indicators at the school level or even at the classroom level. Generally, and especially in the case of more autonomous schools and decentralised education policy, the education process at the school and classroom level is the responsibility of the school. If process indicators at that level are taken into account, however, they should be based on an empirically validated model of education. At the local level or the school level, local policy-makers and practitioners might be more interested in a further refinement of the process aspects of such a model at the school and classroom level.

A model of educational effectiveness that tries to explain the differences in education outcomes between students can serve as a basis for the development of an indicator system, providing at the policy level important process indicators for educational decision-making that are of necessity broadly defined on the one hand and, on the other, provide practitioners and local policy-makers with more precise information about the functioning of the education system at the school level and the classroom level for their own efforts to improve education at these levels (Scheerens, 1990).

The development of an international system of education indicators can profit from national and international studies on school effectiveness or educational effectiveness models. In the following, a description will be given of the International School Effectiveness Research Programme (ISERP), which addresses the question: what factors that can explain effectiveness are generalisable across countries, and what is the influence of

the context on these factors? In the long run, this study can contribute to the development of process indicators on school functioning. The ISERP project starts from an explicit model for educational effectiveness including factors at the levels of the individual student, the classroom, the school and the context.

2. Background

The history of school effectiveness research can be described in different ways. A favourite way is to look at school effectiveness as a reaction to the pessimistic views on teachers, schools and education which greeted the disappointing results of research. Another, quite different interpretation of the history of school effectiveness research considers this research as a natural prolongation of research from the past which dealt with teaching, instruction, curriculum, school organisation, and so on. Depending on one's views of history, different godfathers for school effectiveness research can be named, such as Coleman *et al.* (1966) or Jencks *et al.* (1972). For most school effectiveness research, the work of Edmonds (1979) and Brookover *et al.* (1979) in the United States and of Rutter *et al.* (1979) in the United Kingdom are important starting points.

Especially in the United States, since Brookover *et al.* (1979) and Edmonds (1979), a great deal of work has been done by researchers relating to the early American studies, but this seldom refers to any British or European literature other than that of Rutter *et al.* (1979) and Mortimore *et al.* (1988). In other countries, although references are made to the earlier works of Brookover *et al.*, (1979), Rutter *et al.* (1979), and Mortimore *et al.* (1988), researchers show a similar ethnocentricity caused by the fact that in all countries research has been looking for factors that could contribute to effectiveness within the particular country. The founding of the International Congress for School Effectiveness and Improvement and a journal associated with that organisation (*School Effectiveness and School Improvement*) in 1988, stimulated the exchange of research and research results in the field of educational effectiveness. The International School Effectiveness Research Programme (ISERP) is one of the examples.

3. The International School Effectiveness Research Programme (ISERP)

The ISERP programme of studies sets out to do a number of things. It aims to build on existing models of good practice in research design, and to avoid the variation in national studies' research designs that limits transferability within and between countries, by using standard measures of inputs, processes and outcomes, and common methods of data analysis and data collection (although each participating country can add to the common core any samples, instruments and analyses that may be appropriate for its own knowledge and/or policy needs).

The programme aims to improve on the quality of the national studies so far undertaken, and on the conceptualisation, operationalisation and measurement that are evident in the existing international studies.

A number of studies have been initiated as part of the overall programme. Firstly, the past three years have seen an extended programme of meetings, the production of

discussion papers, and presentations to many academic conferences, as we have sought to make sense of the school effectiveness literature, the international school/instructional effectiveness studies and the varying contexts, both socio-cultural and educational, of the different countries involved. Publications based on this phase have already appeared (Creemers *et al.*, 1991; Reynolds *et al.*, 1994).

Secondly, there is the present pilot study which attempts to develop hypotheses concerning the variables at different levels (class, school, local context, societal context) which have an impact on student learning. For this purpose, "archetypal" schools have been selected which fall at the extremes of the expected distribution, or indeed outside it.

The third part of the ISERP programme is designed to increase our knowledge of effective instructional and school processes by linking together existing projects, including those of the IEA such as the Third Mathematics and Science Study, that are taking place in different countries. This may involve trying to ensure that different projects share a common core of methodology and instrumentation, or alternatively may involve pooled analysis as appropriate. The aim is to make progress in the elucidation of those factors responsible for variation in education system performance.

4. A Study of "Archetypal" Schools in a Range of Countries

The countries/territories involved in this study are Australia, Canada, Hong Kong, Ireland, the Netherlands, Norway, Taiwan, the United Kingdom, and the United States. (Although not every one of these entities is an independent country, each is being treated as such for research purposes, and hence the term "country" will be used below.)

Certain characteristics and principles, given our review and assessment of existing knowledge outlined above, are axiomatic in the generation of our study design:

- Data are being collected on individual pupils, on their classroom processes, on school processes, on the classroom-school interface, and on the meso and macro factors that may promote or hinder student learning.
- These data are on cohorts of students to whom educational processes will be giving varying degrees of "value added" educational and social development, with data being collected on students at intake and at outcome.
- The outcomes of educational processes are seen as both academic and social in character.
- Process variables are particularly intensive at the level of the instructional processes within classrooms, where learning takes place.
- Outside-school factors at macro and meso levels are being studied in detail, since they provide the educational and cultural contexts within which learning takes place.
- Data collection is determined by the need to test models (and generate models) of school and classroom influences upon student learning and development.

In addition to those aspects of the study design that have been determined by our assessment of the principles which should guide us, more pragmatic considerations also determined that:

- We should study younger rather than older children, since the older children become, the greater is the chance of school effects being hidden by being introduced into analyses as intake effects. Also, the younger children are, the greater is the chance of their having no prior learning in the areas of knowledge that we are interested in.
- We should maximise the variance in the dependent variables concerning student outcomes by choosing countries where there are apparently very high levels of pupil attainment (e.g. Taiwan) and also countries where there are apparently lower levels.
- We should maximise the variance in outcomes within societies by choosing "archetypal" schools where there are exceptionally high and low levels of "value added" by the school, besides choosing "typical" or "average" schools.
- We should study educational effects upon areas of the curriculum and knowledge where there is evidence of greatest school effects (e.g. science or mathematics), thus minimising the problems of controlling for home and outside-school influences.
- We should choose as our focus a school subject (mathematics) that is the most uniform internationally in terms of knowledge content, to reduce the problems of measurement and comparability entailed in subjects where different societies teach different bodies of knowledge. However interesting it may be to examine the role of schools in the differential transmission of knowledge through such areas as the humanities curriculum, research needs are for the definitions of subject content to be as near universal as possible across countries. This approach is, of course, beset with definite limitations.

In the design of the research programme, several models of educational effectiveness were compared. According to Ralph and Fennessey (1983), it is important not only to find out what factors make a difference between effective and ineffective schools, which could lead to a fishing expedition for all kinds of variables (see, for example, Levine and Lezotte, 1990), but also to develop models that make clear distinctions between context, input, process and outcomes and between the different levels in the education system too, such as context, school, classroom and student. On the basis of the general model developed earlier by Scheerens and Creemers (1989), both authors developed models for school effectiveness (Scheerens, 1992) and educational effectiveness (Creemers, 1994) in which they try to combine instructional effectiveness at the classroom level, the conditions at the school level and the contextual level. A comparable model is the QAIT model by Stringfield and Slavin (1992), in which a distinction is made between quality, appropriateness, incentives and time. Although these models differ with respect to the number of variables and factors included, and with respect to the emphasis on the individual level, the classroom level or the school level, they have in common that they make a distinction between time and/or opportunity to learn and the instructional quality of schooling. These broad categories are elaborated at each of the levels within the education system (see Figure 5.1).

Such models are being tested in the educational effectiveness research programme within ISERP, and in several national studies.

Figure 5.1. A comprehensive model of educational effectiveness
Modèle global d'efficacité de l'enseignement

Levels	Components/characteristics of quality, time and opportunity	Formal criteria	Criteria for effectiveness
Context	Quality - Policy focusing on effectiveness - Indicator system/policy on evaluation/national testing system - Training and support system - Funding based on outcomes Time - National guidelines for time schedules - Supervision of time schedules Opportunity - National guidelines for curriculum	Consistency Constancy Control	
School	Quality/educational - Rules and agreements about classroom instruction - Evaluation policy/evaluation system Quality/organizational - Policy on intervention, supervision, professionalization - School culture inducing effectiveness Time - Time schedule - Rules and agreements about time use - Orderly and quiet atmosphere Opportunity - School curriculum - Consensus about mission - Rules and agreements about how to implement the school curriculum	Consistency Cohesion Constancy Control	
Classroom	Quality of instruction curriculum - Explicitness and ordering of goals and content - Structure and clarity of content - Advance organizers - Evaluation - Feedback - Corrective instruction Grouping procedures - Mastery learning - Ability grouping - Co-operative learning highly dependent on - Differentiated material - Evaluation - Feedback - Corrective instruction Teacher behaviour - Management/orderly and quiet atmosphere - Homework - High expectations - Clear goal setting - Restricted set of goals - Emphasis on basic skills - Emphasis on cognitive learning and transfer - Structuring the content - Ordering of goals and content - Advance organizers - Prior knowledge - Clarity of presentation - Questioning - Immediate exercise - Evaluation - Feedback - Corrective instruction	Consistency	
	- Time for learning - Opportunity to learn		
Student	- Time on task - Opportunities used Motivation Aptitudes Social background		- Basic skills - Higher order skills - Metacognitive skills

Source : OECD.

5. Research Questions and Design of the Study

Generally speaking, the research questions can be phrased as follows:

- What factors are associated with the academic and social outcomes of students within a particular cultural context, within each particular society, within each country (e.g. within a low social economic status – SES – context)?
- What factors are in general associated with the academic and social outcomes of students across countries, and what factors are restricted to a specific cultural context (e.g. within a low SES context across countries, etc.)?
- What factors are associated across countries with the academic and social outcomes of students with different characteristics?

More precisely, the research questions are as follows:

- What are the differences between schools displaying low effectiveness, average effectiveness and high effectiveness in different countries?
- What are the differences between schools in low SES and middle SES areas within different countries?
- What are the differences between schools in low SES and middle SES areas between countries?
- What factors contribute to effectiveness within countries?
- What are the contributions of the different levels in the education system to educational outcomes in different countries?
- What are the differences between schools displaying low effectiveness, average effectiveness and high effectiveness across countries?
- What factors contribute to differences in effectiveness between countries?
- What are the contributions of the different levels in the education system to differences in educational outcomes between countries?
- What are the contextual factors that can explain the differences between countries?

The study consists of a more quantitative and a more qualitative part. The qualitative part includes case studies within countries and a comparison of these case studies between countries. The quantitative study has a design that could be used for a larger longitudinal survey, but is now being applied to a small sample of schools. In each country, highly effective, averagely effective – the so-called “typical” schools –, and ineffective schools are involved in this study. Within each group of schools a distinction is made between low SES and middle SES schools. From each of these categories at least one school, including two classes, is involved. In some countries, such as the Netherlands, where schools normally do not have two classes per grade, just one class is included. Of course, this makes a comparison with respect to a distinction between class and school factors more difficult.

The selection of schools was based on an analysis of existing databases, when available. Otherwise, schools were selected on the basis of their reputation by those key informers in the education system who are likely to recognise “archetypal” schools which lie at the positive, and negative extremes of performance, or in the centre of the average range. This can always be checked in the analysis of the data later on. The data

on progress in learning, *i.e.* "value added", will be used to validate the classification of the schools into one of the three effectiveness categories.

The data collection takes place at the individual student level with respect to student background, ability, and achievement in both academic and social areas; at the classroom level with respect to teacher behaviour, the curriculum and the organisation of the class; at the school level with respect to the curriculum, the organisation of the school and the resources; and finally at the contextual level. The collection of data started in the school year 1992/93 and will be continued during the school year 1993/94, including research into the transition from one grade level to another. To make international comparison not too difficult, mathematics was chosen as an academic outcome measure. At the start, the children were seven years old, which implies that in most of the countries they were in the second grade of primary school where there is considerable emphasis on mathematics. Between countries, however, there are slight differences, for example in Norway where children start primary education only at the age of seven.

6. First Results

The intake data were collected on mathematics attainment, non-verbal intelligence, occupation and educational qualifications of father/male caretaker, occupation and educational qualifications of mother/female caretaker, and ethnicity. Student outcomes at the end of the school year also included mathematics attainment in addition to other outcome measures, such as locus of control, attendance rate, language ability (a different test for each country), and attitude towards school, self, and mathematics. Table 5.1 presents reliability estimates for CTBS (Comprehensive Test of Basic Skills) and the IQ test taken in the first year. All coefficients were based on all students who took the test and for whom data at the item level were available. The given reliability coefficient is Cronbach's alpha; for the CTBS test this is the reliability based on the raw scores. Because these raw scores are transformed into scale scores, the estimates can also be used as indicators for scale score reliability. The CTBS consists of two parts: computations

Table/Tableau 5.1.
Reliability of the instruments
Fiabilité des instruments

Test name	No. of items	Reliability	No. of students
CTBS 11 comp.	28	0.92	562
CTBS 11 appl.	34	0.88	562
CTBS 12 comp.	30	0.91	2 127
CTBS 12 appl.	42	0.87	2 127
CTBS 13 comp.	38	0.84	449
CTBS 13 appl.	42	0.76	449
IQ part 1	20	0.76	2 526
IQ part 2	20	0.77	2 526
IQ total score	40	0.83	2 526

and application of mathematical concepts. The IQ test also consists of two parts: composition and exclusion.

In Table 5.2, the average scores across schools for the various countries are presented. It should be noted that, because of the limited number of schools and the way they were selected, these average scores are indicative at best; they cannot be considered country means. The means are displayed and the standard deviations are given in brackets. The first CTBS score provides the intake results, and the second the outcome measure at the end of the school year. The levels of CTBS indicate a higher difficulty level. The United States did not use the CTBS test. They used the CAT test instead. Therefore the raw scores on this test cannot be compared to the raw scores of other countries, but a transformation from CAT results into CTBS results is possible. The intake scores for Canada are based on one school. The outcome data for Australia and Taiwan are not available at the moment.

As expected, the Hong Kong and Taiwan schools show high results at intake. The results of the United Kingdom, where mathematics education starts at the age of five, are lower than in Hong Kong and Taiwan and the standard deviation is also slightly higher. The other countries took a lower level of the CTBS as an intake measure. This could be expected, especially for Norway, where, as already mentioned, school education starts at the age of seven. According to the CTBS, the raw scores at the various test levels can be

Table/Tableau 5.2.
Intake and outcome scores for 7 countries*
Notes obtenues aux tests à l'entrée et à la sortie pour 7 pays

	CTBS 11		CTBS 12		CTBS 13	
	Comput.	Applic.	Comput.	Applic.	Comput.	Applic.
United Kingdom	-	-	15.5 (6.4)	23.5 (8.0)	17.3 (8.7)	20.2 (9.8)
Taiwan	-	-	25.9 (4.6)	32.9 (5.1)	-	-
Norway	15.9 (6.4)	21.9 (6.5)	15.8 (5.9)	24.4 (8.0)	-	-
Hong Kong	-	-	26.6 (3.5)	31.5 (5.3)	26.2 (5.9)	27.6 (6.8)
Canada	16.7 (6.2)	24.2 (7.6)	18.1 (5.1)	31.0 (5.9)	-	-
Australia	-	-	17.8 (4.7)	28.7 (6.4)	-	-
Netherlands	22.6 (4.7)	24.8 (4.1)	21.9 (5.4)	25.5 (6.5)	-	-

Standard deviations in brackets.

CTBS = Comprehensive Test of Basic Skills.

* No CTBS scores were available for the United States.

transformed into one scale: the CTBS scale scores. These are presented in Table 5.3. Again, the average scores across schools per country and standard deviation are displayed. The scale scores for the United States were computed by means of the transformation of the CAT scale scores into the CTBS scale scores. As expected, on the basis of the raw scores, the scale scores for Hong Kong and Taiwan are the highest, although in the case of Hong Kong the scale score for application in CTBS 13 is lower than in CTBS 12. The progression from one CTBS level to another probably causes problems. In addition to this, it can be questioned whether the CTBS does reflect what happens in mathematics education in different countries. Countries may emphasise aspects of mathematics education other than those measured by the CTBS. A further analysis of CTBS and mathematics education in the different countries is being carried out. The opportunity to learn – for example, the topics covered by the content of textbooks and the emphasis teachers put on topics in mathematics – is being measured, and the results will be compared with the CTBS scores. In general, scale scores reflect what has already been indicated in the raw scores with respect to the means and the standard deviation.

Because of the fact that the transformation of raw scores into scale scores has caused problems, as the Hong Kong data show, we have refrained from computing progress scores on each test in order to calculate the learning effects for the grade level. First of all, a further analysis on CTBS and the transformation from raw scores into scale scores should be carried out.

Table/Tableau 5.3.

Scaled intake and outcome average scores of 8 countries

Notes moyennes obtenues aux tests à l'entrée et à la sortie pour 8 pays

	CTBS 11		CTBS 12		CTBS 13	
	Comput.	Applic.	Comput.	Applic.	Comput.	Applic.
United Kingdom	–	–	504 (92)	572 (79)	578 (87)	612 (90)
Taiwan	–	–	629 (56)	657 (49)	–	–
Norway	420 (109)	512 (92)	512 (80)	581 (74)	–	–
Hong Kong	–	–	637 (42)	644 (48)	652 (41)	638 (43)
Canada	439 (83)	544 (116)	539 (64)	621 (70)	–	–
Australia	–	–	540 (56)	619 (57)	–	–
Netherlands	517 (71)	552 (53)	585 (61)	592 (57)	–	–
United States	506 (85)	556 (101)	560 (62)	617 (73)	–	–

Standard deviations in brackets.

In order to provide some insights into the effects of background variables, for each country the correlations between information on a student's background and CTBS scale scores were calculated. The results are given in Table 5.4. The information is provided per country, because computing the effects for all students might create artefacts because of the different numbers of low SES students per country and also because the SES categories may have different meanings in different countries. The empty cells in Table 5.4 imply that no information was available on those items. In the samples of Hong Kong and Taiwan there are probably no children who can be described as belonging to an ethnic minority. As can be seen in this table, there is a relationship between SES and the CTBS scores, but generally this is a rather weak relationship. The correlation is the lowest in Hong Kong, which may be an indication of the strong support for school education by parents of all social groups. The "medium size" correlation between IQ and CTBS is an indication of the fact that IQ and aptitude in mathematics are different abilities. As could be expected, there is a negative relationship between ethnicity and the CTBS intake.

Table/Tableau 5.4.

Correlations between student background variables and CTBS scores
Corrélations entre les variables relatives à la situation des étudiants et les résultats au CTBS

	Variable	Intake comp.	Intake appl.	Outcome comp.	Outcome Appl.
United States	SES mother	0.23	0.26	0.03	0.00
	SES father	0.25	0.29	-0.10	-0.14
	Ethnicity	-0.41	-0.53	-0.02	-0.00
	IQ	0.60	0.59	0.29	0.42
United Kingdom	IQ	0.49	0.58	0.45	0.41
Taiwan	IQ	0.36	0.54		
Norway	SES mother	0.10	0.02	0.07	0.18
	SES Father	0.13	0.05	0.16	0.07
	Ethnicity	-0.30	-0.30	-0.34	-0.29
	IQ	0.38	0.36	0.44	0.50
Hong Kong	SES mother	0.05	0.12	0.14	0.19
	SES father	-0.07	0.03	0.07	0.08
	IQ	0.36	0.54	0.43	0.45
Australia	IQ	0.46	0.51		
Netherlands	SES mother	0.23	0.27	0.23	0.15
	SES father	0.11	0.17	0.05	0.09
	Ethnicity	-0.29	-0.23	-0.12	-0.20
	IQ	0.45	0.57	0.36	0.49

CTBS: Comprehensive Test of Basic Skills.

SES: Social Economic Status.

The correlations between background variables and CTBS outcomes are mostly smaller than the correlations between background variables and CTBS intake scores. This might indicate that the influence of background variables is declining over time in favour of the influence of education in schools, although this effect could also be caused by differences in reliability between CTBS 11 and CTBS 12/13.

To carry out a quantitative analysis making use of the HLM multilevel analysis is problematic, because of the fact that in general the numbers of schools and classes are quite low, and in some countries intolerably low. Nevertheless, to get an indication of how the variation between the classroom and the school level between and within countries could be split, such an analysis was carried out. The results of this analysis are surprising: although "archetypes" were chosen – schools in different effectiveness categories – in some countries the between-school variance was zero. But the interpretation has to be carried out with caution, because this could be an effect of the determination of the school model by the two axes of a small number of schools and a small number of classes. Because of the small number of schools, the variance at the school level cannot be estimated in a reliable manner. However, the estimates can give an impression of the variance within each country's sample. As such, the estimates can serve as a basis for the comparison of the various factors found in more qualitative material. For example, when there is no variance at the school level after correcting for intake in a given year, it is difficult to see why a school factor should have a substantial effect. These estimates, however, are less useful for within-country generalisation of the magnitude of the influence of all three levels. The main reason why they cannot be used for this purpose is that the schools were selected because they fell at the extremes or even off the expected distribution, so as to maximise variance at the school level. This would mean a biased estimate of true school-level variance in a country.

Table 5.5 presents the estimation of variance components for various countries for the computation part of the CTBS. An analysis was carried out on the CTBS raw scale scores. The left-hand side of the table gives the estimates of variance with no predictors included, that is, the *total* between-school and classroom variance; the right-hand side of the table gives the estimates of the model with the CTBS computation of intake as predictor at the student level, that is, the "value-added" ("net") between-school and between-classroom variance. In this way one can compare what proportion of the variance and outcome scores was also present at the intake. Table 5.5 presents the estimated variance and in brackets the number of units on which each estimate was based. As can be seen, especially at the school level, the number of units was intolerably small. The initial estimates of the lack of variance for some countries might be due to the small number of schools. The adding of a predictor at the student level in this table explains variance at both the school and the class level. This is indicative of a contextual effect of achievement at class and school level.

It is interesting to see that in some countries the variance seems to cluster at the school level and in some countries at the class level. No class and school-level predictors are included in the model, and this can probably not be done in future, because of the small number of classes and schools in each sample. In the Netherlands, most schools have just one class, so that school and class are confounded with each other. The results seem to indicate that the class level is more important than the school level, which goes in

Table/Tableau 5.5.
Estimation of variance components
Valeurs estimées des composantes de la variance

	Empty model			Model intake predictor		
	Student variance	Class variance	School variance	Student variance	Class variance	School variance
United States	2 437 (158)	460 (9)	0 (3)	1 950 (158)	355 (9)	00 (3)
United Kingdom	6 243 (458)	197 (24)	1 172 (12)	4 532 (458)	0 (24)	423 (12)
Norway	5 789 (399)	26 (19)	452 (9)	4 590 (399)	0 (19)	163 (9)
Hong Kong	1 490 (435)	210 (12)	0 (6)	1 175 (435)	141 (12)	0 (6)
Netherlands	2 883 (134)	141 (8)	385 (6)	2 003 (134)	130 (8)	0 (6)

Number of units, pupils, classes and school scores shown in brackets.

the same direction as the analysis carried out by Hill *et al.* (1993) and Scheerens *et al.* (1989). However, the number of units in our study is too small to take these findings as further support for the results of earlier research.

Yet, we take this tentative analysis as an indication that we have to be careful with factors at the school level and even more so with respect to the contextual level.

Although the qualitative study can bring forth possible interesting variables between schools and context with respect to educational effectiveness, the contribution of these variables to the explanation of the differences in student outcomes seems to be quite small. In order to find the amount of variance that would be significant enough to be explained by these variables, a larger sample of schools, classes, and perhaps also countries, will be needed.

7. Conclusion

The development of a system of education indicators at the school level will probably be restricted to a small set of indicators, such as time and opportunity, and some indicators on the quality of school functioning at the school level and the classroom level. It is important that these indicators should be tested within a model of educational effectiveness by empirical research in national and international studies. International research is being carried out in projects such as those of the IEA, IAEP, and ISERP. ISERP, in particular, sets out to find factors at the school level and the classroom level that do or do not travel between countries, because of the contextual differences between countries. International projects face a large number of problems with respect to the

comparability of outcomes, content coverage, teaching behaviour and more formal characteristics of the education system, such as cohesion and constancy. So far, time for learning seems to be a relatively easy indicator. Opportunity to learn/content coverage causes more problems. And quality of education, which refers to all kinds of variables within schools and classrooms, will be difficult to transform into an internationally comparable, reliable education indicator, although this type of factor can be useful in the national context. In conclusion, this chapter indicates that research and development have a long way to go. Will the sponsors wait?

References

- BROOKOVER, W.B., BEADY, C., FLOOD, P., and SCHWEITZER, J. (1979), *School Systems and Student Achievement: Schools Make a Difference*, Praeger, New York.
- COLEMAN, J.S., CAMPBELL, E.Q., HOBSON, C.F., MCPARTLAND, J., MOOD, A.M., WEINFELD, F.D., and YORK, R.L. (1966), *Equality of Educational Opportunity*, US Government Printing Office, Washington, DC.
- CREEMERS, B.P.M. (1994), *The Effective Classroom*, Cassell, London.
- CREEMERS, B.P.M., REYNOLDS, D., SCHAFFER, E.C., STRINGFIELD, S., and TEDDLIE, C. (1991), *International School Effects Research*, National Kaohsiung Normal University Press, Kaohsiung.
- EDMONDS, R.R. (1979), "Effective schools for the urban poor", *Educational Leadership*, Vol. 37 (1), pp. 15-27.
- HILL, P.W., HOLMES-SMITH, P., and ROWE, K.J. (1993), *School and Teacher Effectiveness in Victoria: Key Findings from Phase 1 of the Victorian Quality Schools Project*, Centre for Applied Educational Research, The University of Melbourne Institute of Education, Melbourne.
- JENCKS, C., SMITH, M., ACLAND, H., BANE, M.J., COHEN, D., GINTIS, H., HEYNS, B., and MICHELSON, S. (1972), *Inequality: a Reassessment of the Effects of Family and Schooling in America*, Basic Books, New York.
- LEVINE, D.U., and LEZOTTE, L.W. (1990), *Unusually Effective Schools: a Review and Analysis of Research and Practice*, National Center for Effective Schools Research and Development, Madison, Wisconsin.
- MORTIMORE, P., SAMMONS, P., STOLL, L., LEWIS, D., and ECOB, R. (1988), *School Matters: The Junior Years*, Open Books, Wells, United Kingdom.
- RALPH, J.H., and FENNESSEY, J. (1983), "Science or reform: some questions about the effective schools model", *Phi Delta Kappan*, Vol. 64 (10), pp. 689-694.
- REYNOLDS, D., CREEMERS, B.P.M., NESSELRODT, P.S., SCHAFFER, E.C., STRINGFIELD, S., and TEDDLIE, C. (eds.) (1994), *Advances in School Effectiveness Research and Practice*, Pergamon Press, Oxford.
- RUTTER, M., MAUGHAN, B., MORTIMORE, P., and OUSTON, J. (1979), *Fifteen Thousand Hours*, Open Books, London.
- SCHEERENS, J. (1990), "School effectiveness research and the development of process indicators of school functioning", *School Effectiveness and School Improvement*, Vol. 1, pp. 61-80.
- SCHEERENS, J. (1992), *Effective Schooling: Research, Theory and Practice*, Cassell, London.

- SCHEERENS, J., and CREEMERS, B.P.M. (1989). "Conceptualizing school effectiveness", *International Journal of Educational Research*, Vol. 13, pp. 691-706.
- SCHEERENS, J., NANNINGA, H.C.R., and PELGRUM, W.J. (1989). "Generalizability of instructional and school effectiveness indicators across nations: preliminary results of a secondary analysis of the IEA second mathematics study", in B.P.M. Creemers *et al.* (eds.), *School Effectiveness and School Improvement. Proceedings of the Second International Congress*, Swets and Zeitlinger, Lisse, pp. 199-209.
- STRINGFIELD, S.C., and SLAVIN, R.E. (1992). "A hierarchical longitudinal model for elementary school effects", in B.P.M. Creemers and G.J. Reezigt (eds.), *Evaluation of Educational Effectiveness*, ICO, Groningen, pp. 35-69.

Chapter/Chapitre 6

The Development of Indicators on Equity in Education
La mise au point d'indicateurs sur l'équité dans l'enseignement

by

Alan Gibson

Guilford, Surrey, United Kingdom

and

Denis Meuret

IUFM de Bretagne, Rennes, France

Information relating to equity, especially on gender issues, is already provided in the various issues of *Education at a Glance* through breakdowns of data in many existing indicators. Network C considers that there is a need to increase the extent to which future indicator sets from the OECD illuminate equity issues. This chapter:

- notes the importance attached to equity by OECD ministerial discussions;
- seeks to set boundaries to the relevance of the concept for the construction of a set of indicators, without attempting the full philosophical discourse which the subject deserves;
- examines the relationship between the concepts of equity and equality, and some of the arguments for and against attempts to achieve equity in educational opportunity;
- provides some questions which educationists may use in order to test their system's commitment to the principle of equity;
- proposes some of the types of indicator which could with advantage be developed in future OECD work.

*

* *

Note de synthèse

L'information relative à l'équité dans l'enseignement, et notamment à l'égalité entre élèves garçons et filles, existe déjà dans les différentes versions de Regards sur l'éducation où sont ventilées les données qui composent un grand nombre d'indicateurs. Le Réseau C a estimé qu'il était nécessaire d'accroître l'éclairage donné par les futures séries d'indicateurs de l'OCDE sur les questions d'équité. Dans ce chapitre:

- on note l'importance attachée à l'équité dans les débats ministériels de l'OCDE;*
- on cherche à fixer les limites de la pertinence d'une théorie de l'élaboration d'une série d'indicateurs, sans aborder le discours philosophique complet que justifie ce sujet;*
- on examine les rapports entre les concepts d'équité et d'égalité et quelques-uns des arguments pour et contre les tentatives visant à instaurer l'équité en matière d'égalité des chances dans l'enseignement;*
- on énumère quelques-unes des questions que les théoriciens de l'enseignement peuvent poser pour voir à quel point leur système est fidèle au principe d'équité;*
- on propose quelques types d'indicateurs qui pourraient utilement être élaborés au cours des travaux ultérieurs de l'OCDE.*

*

* *

1. Introduction

This chapter raises questions about future indicators within the INES project and its possible successor operations within the OECD. Questions on equity in education can only be addressed by a set of indicators derived from an agreed strategy and a co-ordinated effort across all existing networks and the technical group, and throughout any future institutional arrangements for indicator production within the OECD; but equity has been of particular concern to Network C, with its responsibility for indicators on the functioning of schools and school systems.

It is through the operations of schools and school systems that equity can be more or less successfully achieved, and/or inequity moderated; but for information on whether and to what extent equity is approached, INES and *Education at a Glance* depend on indicators which are being developed by other networks and the technical group, for example indicators concerning resources and their distribution, students and their backgrounds, system goals and results for different groups, and public and professional opinions about equity.

2. The Concern for Equity

Equity is a natural concern of countries whose membership of the OECD is founded on a common set of basic values which include "respect for human rights and (...) pluralist democracy" (OECD, 1993a). The specific interest of OECD Members in equity in education was stressed in November 1990, when Ministers of Education of OECD countries met to discuss "High Quality Education and Training for All" (OECD, 1990). In the *communiqué* after the meeting they stated that: "Education for all implies priority for the educationally underserved: to extend the benefits of education to all makes economic sense as much as it accords with social and educational equity..." (OECD, 1993b).

In accordance with these values, certain indicators relating to equity are already provided or proposed in *Education at a Glance*, and Network C has discussed on several occasions how additional education indicators may be developed in order to ensure that equity issues are addressed alongside those of economy, efficiency and effectiveness, quality and standards, with which they are frequently given equal priority.

The nature of concern about equity in education in OECD countries has naturally varied over 40 or more years. Immediately after the Second World War, concern was for equality of access, and the issue of equity was a component of the social discussion of how to guarantee equal opportunities for citizens in accordance with the principles of a democratic society. Next came the "enrolment explosion", the massive increase in post-primary education, accompanied during the early 1960s by extensions of compulsory schooling to age 16. These changes reinforced the role of the school as the main determinant of most young people's future socio-economic status. Thus, by the end of the 1970s, the subject of unequal educational opportunities had become less of a priority because the "democratisation" of access to extended education was reflected in a very sharp rise in the numbers taking full-length secondary education.

In contemporary discourse, as ever more families and children come to see education as a social investment, and qualifications as a means of escaping unemployment, education becomes increasingly competitive. Even for those who belong to the ethnic or social majority, or other majority group in the population, the inequality gap appears to be widening. There is thus a risk of exclusion of substantial groups of people who cannot obtain steady employment and cannot achieve the educational qualifications which would allow them to perceive themselves, or to be perceived by others, as integrated economically and culturally into the rest of society.

Current concerns are therefore about extending to an even greater proportion of the population, even a majority, the standards achieved by the most successful, while at the same time attending to the continuing disadvantages of various minorities. Extending to many more the success of the most successful is seen as a priority particularly in economic terms. The trend away from manual and less-skilled work reduces employment chances for those with low educational attainment, and with skills that are little in demand. It is not enough simply to advocate more education and training of any kind, no matter what, on the assumption, for example, that clerical, service and knowledge-intensive occupations have an indefinite capacity for expansion.

Thus, on the one hand, many of the current issues concerning equity are perennial questions relating to the progress of those from potentially underprivileged groups, whether socio-economic groups, racial and/or religious and/or cultural groups, or groups in particular geographical areas such as inner cities or sparsely populated rural areas. However, on the other hand, contemporary concerns about equity have most recently focused equally on quality, that is, on ensuring not just access for the more or most disadvantaged but, as wider participation has been increasingly established, on access to provision of perceived and attainable value. The questions now, therefore, are not simply about equality of access for the previously "underserved" in terms of removing barriers to their entry in due numbers or proportions. They are, for example, about equal readiness to benefit and about access to equality of worth, and especially economic worth, in the educational experiences to which equal opportunity affords access.

The pursuit of equity is seen to require not only that all have equal access to initial and extended education, but that all are equally enabled to benefit, for example from primary education through pre-primary experience; and that access for underprivileged groups is to educational programmes equal in "worth" to the "worth" of programmes entered by privileged groups. Four examples, cited in *Access, Participation and Equity* (OECD, 1993b), and referring to three different stages of the education system, provide graphic illustrations:

- "It is revealing that when the extent of progress towards gender equality is challenged, then the still-low female participation in technological programmes is stressed as evidence of exclusion from the most desirable educational routes. When, instead, the social composition of access and participation becomes the focus, the comparisons tend to shift to the 'high prestige' programmes as constituting the most desirable.
- (...) some countries have succeeded in recent years in having middle-class students in certain vocational streams and a higher proportion of working-class students in general education (...): [this] may be viewed as a positive development. However (...) quite often privileged groups tend to concentrate on those streams which open up better opportunities for employment, whereas many of the less privileged groups enrol in general education as a second choice, because of lack of places in the more rewarding selective vocational options.
- (...) secondary general-school leavers have to compete for the more attractive apprenticeships, especially in commercial vocations, with intermediate and even grammar-school leavers (...). *The chances of secondary general-school leavers have been reduced to a small spectrum of less attractive vocations (...).* For the more attractive (...) especially those of commercial clerks in industry and banking, often the grammar-school leaver is preferred; but also for the attractive craftsmen's and industrial workers' jobs, the intermediate schools' examination has become widely necessary.
- The secondary-school leaving qualification is necessary for access to higher education, but less and less is it sufficient to provide the candidate with a *free choice of establishment or subject (...)* increasingly depends on (...) additional requirements [such as] entrance examinations, interviews, subjects studied, school

records (...) [and] explicit and implicit hierarchies based on academic prestige are appearing between establishments (OECD, 1993b)."

The issue which has been debated by Network C is therefore not *whether* a set of indicators such as that which is being built up by INES should include more indicators related to equity, but *how* to prepare a rationale so that the *set* of indicators can have a broader relevance to equity, commensurate with the policies and values of OECD countries.

3. Equity and Equality

Equity in education systems is not, of course, a synonym for equality of outcomes, uniformity of provision or standardisation of experience. Theoretically speaking, the achievement of equity in education systems depends, as in other systems, on the "fair" distribution of a certain number of "goods" to various categories of recipient. The equity of education systems is most often judged according to the criterion of equality of opportunity; the use of this criterion does not of course presuppose that every student will succeed equally well, rather that any success will depend in no way on the "hazards of birth" – that is, social category, place of birth, degree of handicap, etc. For a set of indicators to illuminate the distributions sought and achieved in any education system or systems, it is necessary to define both the "goods" and the categories of recipient to be included:

- Among the "goods" could be the fact of being educated for a certain length of time, or of reaching a certain level in an education system, or of attaining certain competences, or mastery in certain fields of knowledge, or certifications. Other "goods" could include the possibility of studying in a pleasant atmosphere, of being treated with humanity and fairness, of having competent teachers.
- The categories for which equity is sought will have some common identities and some different ones across Member countries. Equity indicators are already included in *Education at a Glance* for relative access/achievement by gender; in some contexts for equity between age-groups, to avoid "age-ism"; and for relative performance of the public, hybrid and private sectors (which may include confessional sectors in some countries). Categories of students, the equity of whose educational experience might need to be considered by policy-makers in future, are socio-economic groups, ethnic/cultural groups and affiliated schools, religious groups and affiliated schools, disabled groups, students for whom the language of instruction is a second language, students in particular kinds of geographical area, migrant groups and travelling groups (see list in Annex 6.1; Annex 6.2 also shows breakdowns of existing indicators).

Of course, it would be unreasonable to expect an education system which is embedded in a society which itself exhibits significant social inequalities to achieve even reasonable equality in educational outcome; the greater the social inequalities, the less easily can equality of educational outcome be achieved. But if OECD education systems, in themselves and through the schools which are their component parts, seek to achieve

equity in their own operation or functioning, so that inherent personal and social inequalities are at least not increased and at best reduced by students' educational experiences, a set of indicators such as that of *Education at a Glance* needs to provide policy-makers with illumination on the systems' success with as many categories as possible of potentially disadvantaged students.

If the promotion of equity is a matter for the functioning of schools and school systems, the enabling climates and guiding frameworks are matters for well-informed policy-making. The questions concerning equity policies have been debated intensively and extensively in OECD countries over many years, and have raised many policy possibilities which have been variously emphasised and modified by different governments facing priorities, interpretations and challenges which change over time.

The issues have been and continue to be emotive, contested and fundamental. Arguments in favour of specific policies to decrease educational inequalities are often that:

- children are not responsible for their parents' cultural background or commitment to schooling;
- a part of educational inequality comes from the education system itself (inequalities of quality resources, of judgements/assessments, of teacher expectations); and
- to reduce inequalities by enhancing the level of less successful pupils is a way of increasing the effectiveness of the system as a whole.

Arguments for limiting the policies directly devoted to reducing inequalities are that reducing inequalities in education:

- would not necessarily result in a more equal society,
- would cost too much, and
- might serve to handicap the more able students.

The tension between these two positions on the search for "equity", means that the following important questions have often to be posed, and in each given context answered, by educational policy-makers:

- Is it equitable for education policies and resources to be used to promote policies of a social character, including equity, often called "social engineering" by those who oppose them, or equitable only for education policies not to impede such policies? Or is it inequitable for education policies not to promote such policies if they are the declared policies and goals of the country?
- Is it true that each system, each school, and each teacher, need not feel that inequality of outcomes is wrong provided, without jeopardising the most able students, that everything possible at reasonable expense has been done to secure equality of opportunity? Or should an "equality of probabilities" be expected, by which systems could show that members of different categories have an equal probability of, say, reaching a given level of certification?
- Is it inequitable if schools attended by greater proportions of disadvantaged children have access to only the same resources, in quantity and in quality (for example, of teachers) as schools attended mainly by advantaged children?

- Is it inequitable, when equity is constrained by cost considerations, if advantaged parents, who can devote more time to their children and their childrens' studies, or can prepare their children to accept more readily the disciplines of study and schooling, can also, through their knowledge of education systems, choose more effective institutions for their children?
- Is it true that an inequality which all known policies were incapable of reducing without prejudicing more able students, increasing costs or putting too much emphasis on academic results, should not be considered an inequity?
- Is it true that if one country could be shown, by international comparison of relative success in equity or equal opportunity policies, to know how to reduce any given inequity, other countries should no longer feel able or obliged to ignore the same inequity?

4. The Development of Indicators on Equity

Improved international indicators, relevant to equity for a broader range of categories of student, will not provide straightforward answers for any country to these complex, philosophical and moral questions. But they will enable each country to make better-informed evaluations of policies on equity, if the comparative data can be generated.

Shortly after the compilation of the tables for *Education at a Glance* (OECD, 1995) is complete, at a time when it will be possible to take stock of the full range of breakdowns already achieved, Network C hopes to take forward its considerations of equity indicators, in order to extend the range of information provided to policy-makers for evaluating issues of equity.

At the time of writing this chapter, it would be premature to do more than speculate about the nature of the strategies which might be proposed. Clearly they will not attempt to encompass equity matters in all countries, for all categories listed in Annex 6.1, and at all ISCED levels. Some selection will need to be made in the light, pre-eminently, of policy priorities and likely data availability. However, it is possible to speculate that relevant indicators will fall into four main types:

- Indicators which show *disparities* across different categories of student demonstrating, for example, whether their actual *participation* in particular programmes and/or ISCED levels is proportional to their numbers among those eligible; or demonstrating whether the *expenditures* on the same provision for different groups is commensurate with their respective sizes or, if not commensurate, whether it differs as a matter of deliberate policy. For example, are there disparities in expenditures so that highly qualified or experienced and therefore more expensive teachers are more frequently found in schools with advantaged students?
- Indicators which show expenditures, participation and outcome/success rates for *special programmes* intended to combat inequities at different ISCED levels. The study of the extent of positive discrimination practices, and programmes specially devised for the disadvantaged student, are of special interest because they can compensate not only for inequalities external to the education system, but can also

reduce internal inequalities (which may also be inequities) and accidental or systemic discrimination.

- Indicators which seek to probe the "quality" issues raised above: for example, access to subjects, courses, programmes of study valued by other, non-disadvantaged or majority or well-informed categories; access to more efficient schools where student evaluation practices and links with parents are well-developed to serve disadvantaged students; access to sophisticated advice when "optional" choices are made, within and after the end of compulsory schooling and on into tertiary education;
- Indicators showing the *level of knowledge* in given subject(s) of sample cohorts of different categories of pupils, especially social categories, at specified ages before the ending of compulsory education; such indicators would show the *outcomes* of any differences in quality, expenditures, effectiveness of schools or teachers and other resources provided for the different categories of pupils.

References

OECD (1990), "Communiqué of Ministerial Meeting on High Quality Education and Training for All", November, Paris.

OECD (1993a), "Communiqué of Ministerial Meeting, Council", June, Paris.

OECD (1993b), *Access, Participation and Equity*, Document DEELSA/ED(93)8, Paris.

Annex 6.1

Variables useful for calculating indicators on equity

The two genders (already used as a variable in some indicators)
Socio-economic groups
Ethnic/cultural groups and affiliated schools
Religious groups and affiliated schools
Disabled groups
Students for whom the language of instruction is a second language
Geographical areas and their schools
Age-groups (in 10 or 15-year divisions above, say, age 29)
Migrant groups
Travelling groups
Enrolment in public, hybrid public/private, and private education

Annex 6.2

Indicators in *Education at a Glance* (OECD, 1993) showing (or perhaps capable of showing) disparity by gender, socio-economic group, race, religion or other categories listed in Annex 6.1

NB: Indicators shown are contained in the 1993 version of *Education at a Glance* (to be complemented by the 1995 version of *Education at a Glance*). Almost all the items listed above could possibly show disparity breakdowns by (additional) categories.

- C1 Educational attainment of population by age-group
- C2 Gender differences in education
- C4 Home and school language
- C5 Labour force participation by level of education, with gender differences
- C6 Unemployment by age-group and gender
- P6/7/8 Expenditure per student, by level and relative to GDP
- P10 Ratio of students to teachers by level, and by public/private sectors
- P11 Participation in formal education at ages 5-29
- P12 Participation in early childhood education
- P13 Participation in secondary education
- P14 Transition to upper secondary and tertiary education by age
- P15-7 Participation in tertiary education by gender and age
- R1-4 Achievement in reading, mathematics and science, by students, schools and gender
- R5-7 Upper secondary, and university graduation, by gender
- R7 Science and engineering degrees, by gender
- R9 Unemployment by level of education, by age and by gender
- R10 Education and earnings, by gender

**The Search for Indicators on the Effective Functioning of Tertiary
Education Institutions**

***La recherche d'indicateurs du fonctionnement efficace
des établissements d'enseignement supérieur***

by

Alan Gibson

Guilford, Surrey, United Kingdom

This chapter acknowledges that Network C's responsibility concerning the functioning of schools will in future, in line with the work of the INES project as a whole, need to cover the functioning of institutions in the tertiary sector as well as the primary and secondary sectors. The reasons lie in the substantial expansion, throughout Member countries, of the provision and costs of tertiary education, and the consequent increasing interest in the evaluation of the sector's effectiveness. This interest is reflected, for example, in the decision of the OECD Education Committee to initiate a thematic review of tertiary education policies. The use of indicators within higher education institutions and within countries' higher education systems has been monitored by the OECD Institutional Management in Higher Education programme, and the extended use of indicators is envisaged for management, self-evaluation and accountability purposes. This chapter anticipates that one source of indicators on the effective functioning of tertiary institutions will be suitably adapted versions of indicators identified by Network C on the effective functioning of schools, and an annex provides seven examples of such adaptations. New indicators, which are specific to tertiary education, are also expected, in response to the growing importance for policy-makers of the effectiveness of institutional functioning in this expanding sector.

*

*

*

Note de synthèse

On reconnaît dans ce chapitre que la responsabilité du Réseau C en matière de fonctionnement de l'école devra à l'avenir, conformément aux travaux du projet INES dans son ensemble, couvrir le fonctionnement des établissements du secteur tertiaire, de même que ceux des secteurs primaire et secondaire. La raison en est l'expansion sensible, dans beaucoup de pays Membres, des prestations et des coûts de l'enseignement supérieur et, par voie de conséquence, l'intérêt croissant que suscite l'évaluation de l'efficacité de ce secteur. Cet intérêt se manifeste notamment par la décision prise par le Comité de l'éducation de l'OCDE de procéder à un examen thématique des politiques de l'enseignement supérieur. L'utilisation des indicateurs dans les établissements d'enseignement supérieur et à l'intérieur des systèmes d'enseignement supérieurs des différents pays a été suivie par le programme de l'OCDE sur la gestion des établissements d'enseignement supérieur, et on envisage d'étendre l'utilisation des indicateurs à des fins de gestion, d'auto-évaluation et de reddition des comptes. On prévoit dans ce chapitre que l'une des sources d'indicateurs du fonctionnement efficace des établissements tertiaires sera une version adaptée des indicateurs identifiés par le Réseau C à propos du fonctionnement efficace des écoles, et une annexe fournit sept exemples de ces adaptations. On s'attend aussi à ce que soient mis au point des indicateurs nouveaux, propres à l'enseignement supérieur, en réponse à l'importance croissante que revêt, pour les décideurs, l'efficacité du fonctionnement des établissements dans ce secteur en expansion.

*

* *

1. Introduction

The publication of two editions of *Education at a Glance* in 1992 and 1993, has demonstrated clearly that the INES project is concerned with all levels of education, but the work of Network C has so far concentrated on schools.

This chapter records that the network acknowledges both an obligation and a will to extend to ISCED levels 5, 6 and 7 the indicators on the functioning of institutions. As INES embarks upon the planning for a new mandate, it is right to consider the ways in which indicators on the functioning of tertiary institutions might strengthen the usefulness of the overall indicator set to policy-makers and others. This chapter sets out to demonstrate the importance and value of such indicators, refers briefly to other relevant work, and outlines examples of some indicators which might be considered by the network.

2. The Value of Indicators on the Functioning of Tertiary Education Institutions

A first question to be asked is whether indicators on the functioning of tertiary institutions would be useful to policy-makers in Member countries. It appears extremely likely that they would be, for several reasons:

- Many countries are seeking to increase the number of students in tertiary education, and must therefore regard it as one important contributor to individual well-being, to social health and to economic prosperity. Greater numbers in most countries qualify for entry to higher and other tertiary education and seek, with the active and prudent encouragement of policy-makers, public, parents and employers, to extend their educational careers while young, or to return in later life. Most countries anticipate major increases in the costs of tertiary education for tuition, student accommodation and support.
- The cost per student year in tertiary education is commonly much greater than the cost per year for younger students in schools, in subventions for teaching and student costs and in costs to individuals; many countries are debating urgently how to contain such costs, and how to distribute the burden between public and private expenditures, a matter recently addressed in detail by the World Bank (1994) report on the funding practices in both developed and less developed economies. All countries clearly have a considerable interest in the efficiency and effectiveness of tertiary education.
- Several countries have recently taken steps to improve their within-country evaluation systems for tertiary and especially higher education, which suggests that there is clearly a climate in which it is necessary, and politically acceptable, to call for evaluation of tertiary education. To cite only four examples:
 - France established the *Comité National d'Évaluation* in 1985 to conduct institutional and disciplinary reviews;
 - The Netherlands has established new protocols for external evaluation and inspectorate monitoring of both evaluations and institutional reports;
 - The United Kingdom has recently established quality assurance mechanisms operated by the national funding sources for all tertiary education, and by the institutions themselves acting corporately;
 - The United States now requires that regional accrediting agencies for universities, under legislation initiated in 1988 "to focus on educational effectiveness", assess student achievement and develop standards in twelve areas, including curricula and student and staff achievement; and requires that states set up bodies to examine local colleges and proprietary schools with high student-loan default rates, developing standards covering items such as student performance, graduation and job placement rates.

The importance of the developments in the tertiary field is so high that the Education Committee of the OECD is initiating a thematic review of tertiary education policies, higher education having already been the subject of separate OECD reviews in ten countries. Work already undertaken towards the thematic review has shown that "half of

each generation will soon benefit (or already benefits) from some kind of tertiary education and training experience". The thematic review will be concerned with "quality, efficiency and relevance" of provision at the initial stage of tertiary-level education and training, and will focus on three major themes: diversity of provision and the priorities for public policy; teaching and learning strategies; and strategy implementation.

Thus, post-compulsory education in general, and higher education in particular, either are or are becoming as significant in the economic and social life of Member countries as are schools. Policy-makers have established in many countries new evaluation systems for all or part of tertiary education, and have a clear interest in exploring the value offered by the effective functioning of institutions as well as the costs and flows of students attending them.

3. Existing Work on Indicators on the Functioning of Tertiary Education Institutions

The evaluation of tertiary institutions, including the development of indicators on tertiary institutions, may not be as widely or immediately acceptable in the educational research community as has been the development of evaluation methods and indicators relevant to schools. There may even be some sensitivity among university staff about any suggestion of applying the same kinds of comparison to their work as are applied to that of school teachers. But the traditions of university autonomy established in some countries have been under siege for many years, motivated by increased requirements for accountability, including matters relating to the efficiency and effectiveness of the process. In such circumstances, it may be beneficial for the INES project to propose, in as balanced and sophisticated a context as the *Education at a Glance* publications, the development of international indicators for the functioning of higher and all tertiary education.

There has already been considerable interest within the OECD in indicators in at least the higher education section of the tertiary sector. This has been within the activities of the OECD's programme on Institutional Management of Higher Education (IMHE). Unlike INES, this programme is not concerned to produce indicators for between-country comparisons, but to study the developments in the uses of indicators within institutions and within countries. The programme has two recent publications which have proved relevant to Network C's interest in developing indicators of tertiary education:

- *The Development of Performance Indicators for Higher Education: a Compendium* (two editions, OECD, 1990 and 1993);
- *Dimensions of Evaluation in Higher Education* (OECD, 1991), in which one of the seven chapters, "Cross-National Comparison and Exchange: Higher Education", by John Harris of Samford University, Birmingham, Alabama, makes numerous references to the potential contribution of INES.

The 1990 edition of this publication contains eleven country statements (from Australia, Austria, Canada, Denmark, Finland, France, Greece, the Netherlands, Norway, Sweden and the United Kingdom). Appendices list the indicators proposed or in use in

higher education in four of the eleven Member countries whose systems were studied: Australia, Finland, the Netherlands and the United Kingdom. The second *Compendium* (OECD, 1993) includes statements from two additional countries, Germany and Ireland, and from all of those included in the first edition except Greece. Appendices list indicators in use or proposed for Australia, Austria, France, Germany and the United Kingdom.

Most of the indicators cited as "in use or proposed" are concerned with costs, with the numbers of staff and students, the output of staff in terms of publications, and the output of institutions in terms of student progression, graduate successes and destinations, but some of the indicators are similar to those in the Network C portfolio on the effective functioning of institutions (e.g. indicators on teachers' qualifications, hours taught, staff-student ratios, quality of life on campus, access to libraries and computing centres).

The comments of two authors in the second *Compendium* are of particular significance for any consideration of tertiary performance indicators by Network C and the INES project as a whole.

In his foreword to the publication, H.R. Kells, the editor, writes:

"(...) interest in performance indicators [arose] at OECD meetings (...) in the 1960s as institutions needed information to allocate resources and make programme choices during a period of expansion. In the 1980s (...) governments in some countries sought mechanisms to make differential judgements, to guide funding choices and to monitor the achievement of national goals in an environment of insufficient resources and as they desired to operate more at a distance from traditional patterns of control.

Some important contextual patterns are affecting the interest in and activity about performance measures. First (...) the processes of decentralisation and ex post control begun in the late 1980s continue to unfold. Planning initiatives with goals to be monitored and the increased use of lump sum budgeting is evident in several of the countries.

The second major pattern evident in the country data is a familiar scenario. The whole issue of performance indicators is beginning to be subsumed under the larger managerial process of quality assurance and control. This is occurring in most of the countries in the sample.

(...) one should certainly see during the 1990s more focus on institutional management and institutionally based or accessible information or decision support systems as governments continue their "arm's length" control of OECD [higher education] institutions. At the same time there will be sustained a noticeable interest by most governments in monitoring against national planning goals through indicators and other related data (OECD, 1993)."

A second author, John L. Davies, in a conceptual review, also emphasises the advantages of performance indicators in self-evaluation and management, and thus their relevance to the general functioning of institutions:

"(...) the implementation of performance indicators is (...) a complex effort. Conceptually, they are a prime characteristic of a self-evaluating organisation which

presumes (...) that [they have] permeated and penetrated institutional life in a purposive and systematic manner.

(...) evaluation is action-oriented (...) the interesting question is what universities actually do (...) evidence from successful experiences indicates that responses of various kinds are needed [in the fields of]: staff development and renewal; resource allocation and re-allocation; curriculum development; faculty recruitment; external relations; administrative renewal".

Harris's chapter in IMHE's *Dimensions of Evaluation* (OECD, 1991), makes frequent reference to the INES project, but its particular interest for Network C lies in the importance which it accords, like the 1993 publication mentioned above, to indicators concerned with the processes of tertiary education. Harris cites:

- A paper in which Dahllöf (Uppsala, chairman of the project group which produced the book) reaches the same conclusion as Deming, Ishikawa and Juran, that it is not by focusing on results alone that quality is improved, but by attention to design and production processes, which in education are the curriculum and instruction;
- The value of "mapping" curricula of higher education courses to test comparability, a task which the European Union's ERASMUS project has begun in the fields of business administration, chemistry, history, mechanical engineering and medicine;
- The value of an analysis of university examinations in a range of subjects, and/or of professional admission examinations, in such fields as those mentioned above.

A separate but important study of higher education assessment, covering many but not all OECD Member countries, also refers to the use of performance indicators. The report *Quality Management and Quality Assurance in European Higher Education* (van Vught and Westerheijden, 1992) does not provide details of performance indicators in use in the EU and EFTA countries, but illustrates the proliferation of performance indicators by reference to 14 other publications on the topic, including the *IMHE Journal* No. 2 of 1992.

It is clear, then, from other work in the OECD and elsewhere, that there is already serious comparative study of the ways in which indicators are being used in the tertiary sector to evaluate and in some cases to compare institutions within countries, and are also being employed as part of the evaluation apparatus for both management and accountability purposes.

4. The Nature of Indicators Needed in Future

Many of the indicators within the responsibility of the other INES networks and the Technical Group already either contain, or could be adapted to contain, data on tertiary and higher education, for example those on expenditure, labour force status, training, transitions, even perhaps the attitudes of the general public.

In the field of responsibility of Network C, however, much new thinking is required in order to extend the evidence on effective functioning beyond the school and into

tertiary education. First, the network will need to reconceptualise the notions of functioning which it has used so far, in order to take account of the particularities of tertiary education, since clearly several aspects of school functioning do not transfer directly to the tertiary field. In particular:

- The size of many tertiary institutions makes it necessary to think of indicators on the functioning of component departments or faculties, not least because of the importance of comparing entities with which students identify;
- The parents' interest, though not necessarily without its effects on student achievement in tertiary education, is clearly exercised in quite different ways from those apparent in and integral to the functioning of schools; and
- In most cases, the effects of the locality of the institution are significantly different in tertiary fields from those found in schools, though the effects of the entry abilities and the socio-economic background of students should not be ignored.

Nonetheless, although additional topics which are of distinctive relevance to the tertiary field will present themselves, it is in general justifiable to hypothesise that some of the features of the effective functioning of schools for younger pupils will transfer with due modification to the effective functioning of tertiary institutions. One reason for confidence in this potential for transferability is that Network C's selection of indicator topics for schools was substantially informed by the research on school effectiveness; and the features which the school effectiveness research reveals to be important in schools are largely similar to the features found to be effective in many human organisations, especially those concerned with the production, construction or transfer of knowledge.

For this reason, it is not implausible to be guided by some of these findings in seeking indicators for tertiary education, provided that appropriate transpositions are made to take account of the realities of tertiary systems. The features found to be important may be said to be, in brief:

- the establishment of clear goals or missions, and the pervasive awareness of those goals;
- strong task-focused leadership, and time spent on task (in education, by staff and students); and
- collegial decision-making, and transmitted belief in the importance of achievement by all.

A number of examples of existing indicator topics, identified as relevant to schools but of potential relevance to the tertiary sector, are briefly outlined in the Annex to this chapter. These examples attempt to show some of the insights which Network C would be able to offer, comparable to those being provided for schools, by transposition of indicators to the tertiary setting. The examples are offered in order to invite response, and inevitably have not yet had the benefit of wider discussion. Between the time of the writing of this chapter and its appearance in print, the examples may well have formed the basis of consideration in the network or a working group within it.

It is perhaps important to note, finally, that the selection of indicators in the Annex is not intended to suggest that indicators which had their origin in discussions on schools are stronger candidates for attention than authentic indicators which are specific to tertiary

education and may be identified in future. Such identification of new indicators is to be expected, once consideration of the effective functioning of tertiary education begins in response to its clearly increasing importance for policy-making in OECD countries.

References

- DAHLÖF, U., HARRIS, J., SHATTOCK, A., and VELD, R. in 't. (1991), *Dimensions of Evaluation: Report of the IMHE Study Group on Evaluation in Higher Education*, Jessica Kingsley, London.
- OECD (1990), *The Development of Performance Indicators for Higher Education: A Compendium for Eleven Countries*, Paris.
- OECD (1991), *Dimensions of Evaluation in Higher Education*, Paris.
- OECD (1993), *The Development of Performance Indicators for Higher Education: A Compendium for Twelve Countries*, Paris.
- VAN VUGHT, F.A. and WESTERHEIJDEN, D.F. (1992), *Quality Management and Quality Assurance in European Higher Education: Methods and Mechanisms*, Centre for Higher Education Policy Studies, University of Twente, Enschede.
- World Bank (1994), *Higher Education: The Lessons of Experience*, The World Bank, Washington, DC.

Annex

Examples of possible indicators of effective functioning in tertiary education (adapted from indicators developed for the school sector)

See individual examples for reference to levels of the system to be used for analysis; and see final note on possible breakdowns.

1. Teacher working time

- a)* In the tertiary sector, this indicator would refer to job specifications, such as working hours per week and amount of time spent on instruction, preparation, correction of students' work, administrative tasks and extra-curricular activities and research;
- b)* It would be important to use appropriate definitions of "teachers", e.g. to distinguish (at least) between tenured/not tenured, head of department, lecturer, graduate assistant and others.

2. Hours of instruction received per student per year

- a)* In the tertiary sector, this indicator would refer to the time when a student is with a teacher in a lecture, demonstration, class, group or alone, working on subjects in the curriculum;
- b)* It would be important to take account, for example, of the minimum amount of time for which a student is expected to be in attendance at the institution and the number of years for which any particular programme is studied out of the total number of years of attendance.

3. Time on task

- a)* In the tertiary sector, this indicator would refer to the average length of time spent on tertiary education work by students during a year, including time in lectures, demonstrations, classes, in groups or alone, in private study, and work in laboratories (including computing and other laboratories, whether supervision is by tutorial staff or others);
- b)* It would be important to take account, as above, of the number of years for which any particular programme is studied out of the total number of years of attendance.

4. Educational leadership of the institution, faculty, etc.

- a)* In the tertiary sector, this indicator would refer to the action taken by management at any level to foster the provision of quality education, and to show the involvement of management in leadership of the education process itself;

- b) The levels of the tertiary system to be considered would be institutions, or any component parts of institutions perceived as entities autonomous enough to take such action, such as subject departments, faculties, course or programme teams;
- c) It would be important to take account, for example, of actions intended: to develop co-operation and mutual help between teachers; to foster the implementation of suggestions to improve teaching; to organise help for new teachers from more senior colleagues; to encourage discussions between teachers on students' progress; and to help teachers to form and express effectively opinions on the operation of the institution, faculty, programme or course.

5. *Staff co-operation in teaching*

- a) In the tertiary sector, this indicator would refer both to the degree of explicit collaborative action among the staff and the more intangible sense of collegiality;
- b) The levels of the tertiary system to be considered would be institutions and, because of their size, perhaps more frequently the component units, such as subject departments, faculties and course or programme teams;
- c) It would be important to take account, for example, of:
 - co-operation: the extent of formal arrangements for meetings on teaching matters for staff of the subject, faculty course or programme, and attendance by those eligible; the involvement of staff in the collaborative planning of teaching methods, and the setting of achievement standards;
 - collegiality: opinions on the extent of agreement with other staff in the subject, faculty, course or programme team; and on the curriculum content, teaching styles, organisational matters, professional development of staff, and welfare of students.

6. *Success-oriented ethos: emphasis on student achievement*

- a) In the tertiary sector, this indicator would refer to a success-oriented ethos which has an impact on the academic achievement of students;
- b) It would be important to take account, for example, of policies which encourage constructive signals from teachers concerning students' capacity for success; notification of students' interim academic achievements; attention given to attendance at lectures, classes, etc., and completion of assignments; enhancement of expectations and performance of low achievers; frequency of intermediate assignments and perceived quality of feedback through marking and comments; quality and quantity of private study; student learning habits including attendance and private study patterns.

7. *Decision-making*

- a) In the tertiary sector, this indicator would refer to the levels within education systems at which the essential operational decisions are taken;
- b) The levels of each system to be considered would be the national authorities or the system itself, the intermediate levels, and the institutional and intra-institutional level such as the faculty or department, and the course or programme level;
- c) This indicator would also refer to the areas or domains of decision-making which are the responsibility of the institutions; and the modes of this decision-making within institutions;

d) It would be important to take account, for example, of decisions on:

- attendance/enrolment; progression to successive stages of a course; periods of instruction required; textbooks or other instructional materials; methods of grouping students for instruction; student welfare support and assistance; teaching methods; assessment regime to be applied to students' work;
- the opening or closure of a programme or course; the range of subjects available in an institution; the definition of the syllabus for each subject taught; the setting of examinations and questions within them;
- hiring and firing of staff, at different grades; terms of service and duties of staff; fixing of salary levels for staff, at different grades; career promotions for staff, at different grades; allocation of resources to and within the institution.

Note

In each of the examples given above, breakdowns could be envisaged, where data permitted, by ISCED levels 5, 6 and 7; by institutions showing disparities, by subjects, faculties and programmes; by region where appropriate; and between public and private institutions.

Chapter/Chapitre 8

**The Use of Performance Indicators in School Improvement
and in Accountability**

*Les indicateurs de performance au service de l'amélioration
et de la transparence de l'école*

by

Carmo Clímaco

Department of Programming and Financial Management,
Ministry of Education, Lisbon, Portugal

The purpose of this chapter is to give an account of the endeavours made in Portugal to develop educational performance indicators to be used in school improvement through self-monitoring and self-evaluation, and in accountability. A model of 50 performance indicators that best describe the complexity of schools and of the education system was initially selected on the basis of three criteria: relevance, utility and meaningfulness. A group of schools was invited to implement self-monitoring systems to observe the processes schools would have to go through in order to improve educational quality. One of the first aims in the study of performance indicators was the organisation and implementation of the school information systems and the development of schools' capacity to use them. Schools were therefore trained in self-evaluation procedures and in school policy-making by using performance indicators. A second purpose of the study was to identify the set of indicators that may be used as a common core in accountability processes and in schools' reporting to their communities. So the initial model of performance indicators was refined to seven factors which could be calculated for the universe of schools and corresponded to the teachers' perceptions of school quality and the education system's productivity.

*
* *

Note de synthèse

Ce chapitre a pour objet de rendre compte des initiatives prises au Portugal pour mettre au point des indicateurs de performance de l'enseignement destinés à améliorer l'école au moyen de l'auto-évaluation et du suivi, tout en contribuant à sa transparence. Un modèle comprenant cinquante indicateurs de performance qui donnent l'image la plus fidèle de la complexité des écoles et du système éducatif a été retenu en fonction de trois critères, à savoir, la pertinence, l'utilité et l'adéquation à la réalité. Un groupe d'établissements a été invité à mettre en application des systèmes d'auto-évaluation pour étudier le cheminement que doit parcourir l'école pour améliorer la qualité de l'enseignement dispensé. L'un des premiers objectifs de cette étude des indicateurs de performance était l'organisation et la mise en œuvre des systèmes d'information des écoles, et le développement de la capacité de s'en servir. Les écoles ont donc été formées aux procédures d'auto-évaluation et de prise de décision au moyen des indicateurs de performance. L'étude avait pour deuxième objectif de définir une série d'indicateurs pouvant servir de tronc commun aux procédures utilisées par les écoles pour rendre des comptes à la collectivité ambiante. Le premier modèle d'indicateurs de performance a été affiné pour comprendre sept facteurs qui peuvent être calculés pour le milieu scolaire et qui correspondent à l'idée que les enseignants se font de la qualité de l'enseignement et de la productivité du système.



1. Introduction

The monitoring of the national education system constitutes one of the four major areas of concern for the educational reform in progress since 1986. Its intentions are to establish the principles and to develop the strategies and mechanisms that will allow for the observation and continuous supervision of the performance of individuals and institutions, and to develop various methods of institutional evaluation and self-evaluation. This concern has led to a review of:

- the national information system, namely the timing, collection and treatment of educational statistics;
- the policies and procedures for assessing students;
- the redefinition of teachers' careers and professional status; and
- the evaluation of schools and the monitoring of their performance.

The production of information on educational performance, through performance indicators, raises issues of policy and method that frame the current debate about the improvement of educational quality – about the concept of quality and the measures that operationalise it; the use of information for accountability purposes and for sustaining the reform of the institutions; the grounded decision-making and the negotiation of interests; the problem of legitimacy in evaluation and the choice of the evaluators; and the balance

between generalisability and specificity of information, so that the features that characterise the national education system and the dimensions of educational management may be taken into account.

The debate on these topics has been intimately related to questions about the description and evaluation of the "objects" of education: the concept of information (Comfort, 1982); the relevance, meaningfulness and utility of information (Blank, 1993; Mayston, 1991); the purposes of evaluation, which would determine the selection of the subjects of evaluation, and the respective models and methods (Richards, 1988); and the appropriate format for the communication of the results (Blank, 1993).

The Quality Debate

The issues of "quality" and "efficiency" have been more clearly addressed over the last decade, but to a certain extent they have been implicit in political discourse for the last twenty years in Portugal. The changes which had occurred in the Portuguese political context and in the social tissue of schools by the end of the 1970s made it difficult to make explicit those concepts and relate them to coherent action.

The early 1970s were characterised by several reformist initiatives, the purpose of which was:

"to guarantee to everyone a longer basic education and to provide progressive equality in the cultural access to the different levels of schooling (...) to guarantee the individualisation and diversification of schooling, following the principle of equality of opportunities, regardless of socio-economic status (Ministerio da Educação Nacional, 1973, p. 14)".

In the programmed reform, the concern with "the need for higher levels of achievement" in all grades of the education system is a consequence of the intended "correlation between education and the social functions to be played in the society" (Ministério da Educação Nacional, 1973, p. 14).

The changes which took place from 1974 onwards were determined by the new social and political concerns. The reform of education was viewed as the condition necessary "to create a pure democratic awareness throughout the nation" and "places the school in the [context of the] global questioning of Portuguese society" (Programa do MFA, 1974, § 8b). This document overtly stresses the political above the educational dimensions involved in "the change of schools", "access to school for all", and "the dynamisation of school life". In various ways, these political dimensions permeated school behaviours and challenged the traditional stability of the Portuguese school: stability in the characteristics of the student population, who predominantly came from the same social strata; stability in the contents and methods of teaching; stability in the recruitment of teachers and in their training.

Different attempts "to benefit from experience", "to distinguish democracy from demagoguery" and "to introduce some essential discipline in the running of school life, to guarantee the functioning of the education system" (Decree-Law 769-A/76) were made. However, only by the end of the decade was it possible to "balance the excesses of the

revolutionary period" and "attempt the introduction of a new educational order" (Carneiro, 1980).

The concern for the democratisation of schooling and for the quantitative expansion of education, which were genuine goals of the Revolution, gave way to concern for the qualitative aspects of the functioning of the system, which were formulated in the preparation of a new Educational Reform and an Educational Medium-Term Plan.

"Quality" is not stated as an explicit concept but is perceived as an underlying issue, as can be seen from the new policies on "equality of opportunities", "school resourcing and staffing", "in-service teacher training", "evaluation of students" and "careers guidance".

Texts on the educational quality of Portuguese schools frequently mention the shortage of resources, the need to rationalise those available, and the importance of evaluating the education system and its programmes. They very often make comprehensive inventories of items, describing general features of the system that could be identified as quality descriptors. They cover the different aspects of teachers' education, curricula and learning materials, teaching and learning approaches, the policy and procedures of evaluating students, technological equipment and funding, school organisation and functioning, and administrative structuring and behaviour (Exame das Políticas Nacionais de Educação, OECD, 1984).

The evaluation of the education system is frequently referred to as a means of providing a basis for the Reform policies and is considered an expression of the concern for quality. Costa (1981) considers that "the absence of the system's evaluation policy has to be understood as a national incapacity" and "a reinforcement of inequality and social reproduction".

At the beginning of the 1980s, the future Minister of Education, Carneiro, in the context of international co-operation, proposed a programme for general educational improvement, in which evaluation is the main concern, appearing as a correlate of quality. That programme covered the following aspects:

- the development of systematic studies and critical evaluation of educational structures and planning mechanisms;
- the improvement of a set of education indicators to measure the results, the quality and the evolution of the system;
- the creation of regional and national services to support research and innovation;
- the training of experts in evaluation and systems analysis.

By the end of the decade, da Cunha (1991) states that "increasing democratisation is viewed as an imperative factor of educational reform". Hence, "democratisation" and "evaluation" of the education system and of schools became two interactive goals in the development of educational policy, enshrined as two components of "quality". While the former is one of the goals of education, together with the development of individuals and of society, the latter becomes the strategy to achieve it. However, the concept of "quality" is never made explicit in operative terms, which may explain the relatively low pressure for achievement and the lack of explicitness of the targets of schooling in strategic terms.

Initiatives for the Monitoring of Quality

Various entities are at present involved in studies on the effectiveness of schools and the quality of their functioning, following a variety of methods and focusing on different levels of the education system.

One of these studies has been conducted at the Department of Programming and Financial Management (DEPGEF). Its purpose is to select the performance indicators that best describe the complexity of schools and of the education system, and to identify the set of indicators that may be used as a common core in accountability processes and in schools' reporting on their performance.

This study may be considered an indirect effect of Network C of the OECD INES project, since the discussion and the work developed at the international level inspired and guided the method and the tools designed at home. The national project was planned as a contribution to the international development of performance indicators on the functioning of schools, but was also viewed as an approach to school improvement. Therefore, parallel to the production of school-level data, schools were trained in self-evaluation procedures and in school policy-making by using performance indicators, as an approach to school improvement. So the first aim of this project was the organisation and implementation of the schools' information systems and the development of their capacity to use them.

In the following pages, a description of the project developed at DEPGEF and of some of its findings will be presented.

The Organisation of School Information Systems

Schools are not used to mobilising for their own benefit all the information which they usually collect and send to the central services to produce national statistics. The only thing they derive from the numbers they collect is the boredom and bureaucracy of the task. When asked about specific information to characterise their schools, teachers either do not know, or only have a "global image".

Frequently, they provide an imprecise description of reality, based on their perceptions of classroom events, which are the ones that affect them directly (McLeod, 1988; Sparkes, 1991). They easily fall into the trivialities of their experience, talk about details, but miss the overall dimensions of their schools.

Teachers do not know how far from, or how close to those of other schools their school's problems and performances are; or to what extent they should set more demanding goals or higher levels of achievement. They do not know how many students, and which ones, "survive" throughout their school life, that is, they do not know what rate of achievement there has been in their school over recent years, and how many students usually achieve in all subjects; how qualified and experienced the teaching force is in each place and what local consequences we can derive from this.

In visits to schools it can frequently be observed that even school managers have difficulty speaking about facts. Rather, they make assumptions on the functioning of their schools, which constitute "different forms of recipe knowledge, tacit understandings and pre-understandings" (Sparkes, 1991, p. 6).

In developing a study on the "quality of the educational institutions through the use of performance indicators", we therefore proposed:

- to design and work out a model of education indicators that might provide a description of the performance of the education system and of its basic organisational units;
- to support schools in organising and developing their own information systems so that they could answer demands for accountability and use these systems in management and as a basis for grounded decision-making;
- to articulate state control mechanisms with self-regulation practices at the school level, so that the meso-systemic dimensions of schools might be recognised, from a perspective of decentralisation and autonomy;
- to develop school performance images or profiles, which might act as "operative referents" for evaluation and comparability purposes; and
- to promote the policy and practice of self-evaluation and reflection at school level.

The implementation of a school information system that would provide better knowledge of students' and teachers' work and working conditions implied the clarification of the criteria that have been suggested to be meaningful and relevant in research literature on school effectiveness for the evaluation and measurement of educational quality. It was also necessary to examine the factors that teachers and administrators perceive as identifiers and descriptors of quality, and hence as useful indicators for the management of the education service, and to consider the tools and mechanisms which could produce more knowledge about schools and increase teachers' and administrators' capacity to interpret that knowledge.

We believed that beyond the problems of method and "empirical evidence" surrounding the issue of indicators, of measures and of measurements in education, stressed by Murnane (1987), the very existence of measurement devices would help to achieve better results and to evaluate the effectiveness of policies and strategies, both at national and local levels (Oakes, 1989). Those issues appear in the debate together with the notion of strategy itself (Fielder and Fugl, 1991), and with a different view of policy-making. As Lagerweij and Voogt (1990) suggest, there is no change in schools without a change in the style of policy-making.

From our point of view, the nature of the change to be introduced implied a change in the cultural attitude towards quality itself and its management (Horowitz, 1990). We assumed that "quality change" is more effective when teachers are involved in the process of reform – in its conception, planning and monitoring. Supported by Scheerens (1990), Simons (1988) and other authors, we postulated that "quality" and "change" have to be learnt in order to be effective and cannot be acquired in isolation. They involve the institution as a whole in its systemic and meso-systemic dimensions, and imply an awareness of the different facets of quality perceived by its different actors. Only a review of the education goals and of the practices developed at each level, based on sound information that may balance facts and perceptions, can sustain a productive awareness.

Three concepts were related to quality change: information, indicator and performance.

The concept of "information" and "information process" was inspired by Willensky (1976), who presents the notion of "information process" as the "intelligence function of an organisation" (cited in Comfort, 1982, p. 126). It implies the collection, processing, interpretation and communication of the political and technical data necessary to support decision-making processes. It is "intelligence" to the extent that it attempts to make phenomena understandable and to give insights which better explain the reality. But it is "intelligence" also because it scrutinises facts, explores information in its different meanings and "gives sense" to the data.

The concept of "indicator" is associated with the economy of information, but goes beyond its statistical expression. It is rather a piece of information pointing to a situation deserving further enquiry (Bernard, 1991), and related to the objectives set previously for action (Scheerens, 1990). When taken into systems, indicators profile patterns of performance and act as a cue for reflection and judgement.

The concept of "performance" leads us beyond the notion of results and outcomes. Being a multi-faceted metaphor, it implies "quality results", or a product of competence and meta-analytical awareness as a condition for improvement. Performance, if we take into consideration the dictionary definitions as well as the meanings we put into the usage of this concept, deals with the "visibilities" and "invisibilities" of action, and covers the results as well as the contingencies of action. By "school performance indicators" is meant critical information in the form of statistical statements, covering aspects such as students' achievement and development, the contextual variables from the social environment, and the school organisation, which affect both the teaching and the learning processes, and the resources and the use that is made of them. Thus "indicators" can have an evaluative function whenever they describe critical features of educational organisations and lead to a judgement of their worth on which to base a decision.

2. Method

A number of procedures were followed for the development of school information processes. A central team located in DEPGF studied the issues related to the development of school performance indicators and to the educational measurements involved. At the same time as the development of the work for Network C of the INES project, they were also involved in the organisation of a school monitoring system and in the production of a system of indicators that could be used in the diagnosis and monitoring of schools' performance and serve the purposes of the evaluation of the education system and of the self-evaluation of the schools.

The reform of the educational statistics system provided a substantial source of information, which could feed a permanent observation of the quality of the education system. On the other hand, the implementation of policies on the various dimensions of school autonomy, made schools receptive to internal readjustment or reorganisation of their information systems and self-evaluation practices.

The Initial Model of Indicators

With the inspiration of Scheerens' (1990) model of school performance indicators, it was possible to organise a model of an information system covering 50 indicators around four descriptive dimensions of schools:

- The *context*, or the students' family background factors and the schools' organisational features that may constitute initial advantages or disadvantages for learning:
 - the school's accessibility,
 - the parents' level of education,
 - the parents' socio-professional status,
 - the level of poverty,
 - the pupil : adult ratio,
 - the pupil : teacher ratio,
 - the pupils per teacher,
 - the pupil : non-teaching staff ratio,
 - the extra-curricular activities,
 - the specific support for pupils' learning,
 - the level of organisational complexity.
- The *resources* allocated to schools or those they can mobilise, and the "use that is made of them". Four aspects were covered: physical, human, technological and financial resources. They include variables such as:
 - the rate of occupation of school facilities,
 - the rate of library utilisation,
 - the quality level and condition of school facilities,
 - the teachers' teaching experience,
 - the teachers' qualifications for teaching,
 - the teachers' involvement in education and training activities,
 - the rate of audio-visual equipment utilisation,
 - the budget invested in pedagogical dynamics,
 - the average expenditure per "achiever".
- The *functioning*, or the processes they follow to move from initial conditions and context to success as places of work and living. Nine aspects were selected:
 - the cohesion and participation in school life,
 - the communication and information in school,
 - the time on task,
 - the absenteeism,
 - the safe atmosphere,
 - the pupils' expectations,
 - the socio-cultural intervention in community life,
 - the level of autonomy,
 - the style of leadership.
- The educational *results* covered indicators related to achievement and satisfaction:
 - the rate of achievement per cycle of studies,
 - the quality of achievement,
 - the average age of achievers,
 - the rate of student transfers,

- the rate of drop-out,
- the topic coverage,
- the level of satisfaction.

In the selection of performance information and performance indicators we were looking for their relevance to the educational research literature, the importance they might have for the decision-making processes of the administration and schools, and their meaning in the teachers' images and concepts. The technical feasibility of the indicators was a further selection criterion and justified why some indicators related to facets commonly accepted as important descriptors of school quality are not shown in the list, or are approached through proxy measures (Oakes, 1989).

The indicators selected are very different in importance and in nature. Some will provide quantitative information, suitable for the description of some aspects of resource management, and meaningful for the economy of the system. Some others are more concerned with the qualitative aspects of the context and functioning that account for the differences in educational results. They involve variables that research has associated with school effectiveness. Both are integrated into a comprehensive model of performance indicators to the extent that they relate to the *visibilities* and to the *invisibilities* of schools, and intentionally combine quantitative and qualitative paradigms as a means of responding to the challenge of the plurality of points of view (Ardoino, 1980) in descriptions of the complexity of schools.

The Selection of Schools

Starting from that selection, we proposed to study which indicators could be "networked" and might constitute a common core in the schools' monitoring process and hence be generalised to every school, or to a meaningful sample of schools, to get an image of the education system's performance. The production of a limited number of performance measures that could act as referents for the evaluation of schools, and respond to their accountability needs, appeared to the team to be a first priority.

From the perspective of strategic planning, it would be important that schools should organise their respective information systems to cover the 50 indicators in the model. However, they might also need to "skip" some or add others, or to "desaggregate" some of those selected, so that the model would fit their local needs for self-monitoring and development. Their information systems might thus always be expanded and improved, depending on local conditions. Which indicators schools selected, which were expanded and added appeared as a second priority in the study.

Twelve schools from different Portuguese regions and from different educational levels were involved in the study and associated with the central team. Some of the schools invited to participate had been suggested by the local educational authorities in response to a request that they recommend schools usually regarded as being willing to become involved in innovative projects. Other institutions were invited to participate in this study also, so that we could work with schools from different zones in the country and of different types. For reasons of accessibility, all schools participating were located

in urban areas although the student population might come from rural, industrial or urban zones.

After agreeing to participate in the project, all schools discussed and tested the instruments which had been developed. In each school, the teachers themselves collected the necessary data and, from their own experience, contributed to the reflective analysis of issues to do with the organisation of an information process, with the interpretation of the meaning of performance indicators and with their use in decision-making and self-evaluation.

This approach was the result of our own experience and knowledge of teachers' professional learning processes and successful staff development policies. It is known that innovation and change processes are more valid whenever schools are taken as the basic units of reform (Simons, 1988), when school staff are involved from the conception of the projects (Bollan and Hopkins, 1988), when projects are designed on the basis of common and shared beliefs (Fullan, 1991, 1992), when a positive level of performance and achievement are chosen as the starting point of innovation and institutional development (Aubégnny, 1989) so that the effects may be perceived as being worth the effort (Crandal *et al.*, 1986), and when local contingencies are taken into account (Scheerens, 1990).

Results of the Initial Survey

As a result of the initial work, two documents were produced. The first is *Roteiro de Informação de Escola* (A Guidebook for Charting School Information), which contains all the information and data relative to the indicators selected. This document was conceived as the information process organiser, assuming that "information" means "data" in general, but it also includes the process of collecting, treating and interpreting data, as well as the dissemination of information obtained, the purposes to be followed and the different audiences to be addressed. The organisation and production of the *Guidebook* was simultaneous with the production of the INES project handbook, and clearly reflects the latter's influence.

This *Guidebook* provided an overview of the monitoring process to be used, introduced the issues of self-monitoring and self-evaluation, and presented the devices necessary to gather the information and obtain the indicators. These devices included a description of each indicator and charts for recording the data by which they were operationalised. In fact, most of this data existed – albeit in dispersed form – in different statistical registers in the schools.

The devices for the qualitative indicators involved a set of questionnaires which were addressed to teachers, pupils, parents, and other adults working in the school. All data could either be handled locally and/or sent to the central team for analysis.

The purpose of these descriptors was to map each indicator within each dimension describing schools, and to establish a common language and common guidelines to guarantee their validity for comparisons. Having in mind the users of this document, its training function was respected. Plain language was used to define each indicator, to state its objective, to develop the respective rationale and operationalisation, and to give the

measurement instructions. Related research information on each indicator was added so that teachers could have access to it, and the distinction between "knowledge" and "opinion" could be made evident. In this way, the development of the teachers' interpretative competence and their evaluative capacity was supported. It was also a means of introducing a different register into their professional discourse, more often than not stereotyped by emotions and corporate attitudes.

Schools were asked to reflect on the capacity of the performance indicators for institutional self-evaluation, meaning that they were asked to reflect on the extent to which the performance indicators were powerful tools for the characterisation of educational quality and could guide the identification of priorities in the development of their educational programmes. For that purpose, group and plenary meetings were held at each school, which contributed to sharing thoughts and reactions.

For most of the schools, these project group meetings were their first experience of being scrutinised from the outside and of being compared with others. Most teachers work in isolation and have a partial vision of their schools and of their work: they only know their students and classes, or some few others. The discussion on the issues relating to the indicators – definition, pertinence, validity, operationalisation and measurement – and later on, data and interpretation, provided the teachers with an overview of their school's performance, made them experience the difficulties of the process of making comparisons or evaluations, and helped them to perceive the meaning and importance of being accountable. It allowed them to gain some distance from the "immediacy" (McLeod, 1988), or immediate needs, of their classroom work or managerial "domestic work" and distinguish between "school policy" and "school plan of activities", between managing schools for compliance and managing for performance development, between being rated against a checklist of indicators, which provides an "atomistic view" of schools and "reifies" educational action, and using a system of indicators that becomes an organiser of items to describe performance.

Before data were collected, it seemed important to know the teachers' attitudes and beliefs, and what hierarchies and linkages they established among the proposed indicators. Teachers were therefore asked to rate on a 10-point scale the importance of the indicators selected for the evaluation of their schools.

The second document is the *Monitorização das Escolas – Indicadores de Desempenho* (Monitoring Schools through Performance Indicators), which is an attempt to present school performance measures in a systemic mode. Within the information process, this report corresponds to the stage of data dissemination, which has a strong informative and formative function. We displayed the information using a chart-essay approach (Jones and Mitchell, 1990) for each indicator. In this way we reviewed the respective rationale, selected basic research information on each topic, and discussed and interpreted the findings. The purpose of this work was to find average measures of performance, which could be used as referents for comparisons. It was also intended to lead teachers to research-based judgements, to set the frontier between personal opinions and data, and to train them in self-criticism and in rigour of information.

As had been the case with the organisation of the information systems and with the design and development of the respective tools, the findings were devolved to those

schools which were associated with the analysis and discussion of data. Thus they used and evaluated the school performance monitoring model and tested its capacity to promote "double-loop" teaching and reflective institutions in the sense in which Argyris and Schön (1978) define them. The devolution of information to schools seemed essential to promote institutional maturity. In this devolution process, the central team acted as a consultant and as initiator of the self-criticism.

The Development of Qualitative Indicators

In the meantime, in order to respond to teachers' suggestions, the necessary materials to study and calculate some qualitative indicators related to school functioning were developed. Sets of questionnaires addressed to teachers, students, non-teaching staff and parents were designed for inclusion in the second collection of data. In each school, a sample of respondents was chosen: two to four classes of a terminal grade (the 6th or the 9th grade, according to each school's situation and choice), their respective teachers and parents, and non-teaching staff working at the school. For practical reasons, the samples had to be small enough to be manageable within the teachers' overload of work and constraints of time and available resources. The purpose was to gain an awareness of existing perceptions and opinions. However, the rigour in sampling had to be subordinated to people's willingness to respond.

A total of 1 050 subjects were involved, though their distribution per questionnaire varied with the number of schools volunteering to collect qualitative data. Thus, while seven schools decided to study "participation in school life", gathering 958 subjects, only six of them responded to "satisfaction", involving 787 subjects. The distribution of each population per questionnaire (Table 8.1) adds a further meaning to the response rate, which in the case of the students, for example, is 53 and 80 per cent respectively.

Revision of the Set of Indicators

After the first collection of data, it was possible to refine the model of performance indicators, by reducing the initial set to the most meaningful and useful. This phase of work was conducted in two steps. In a first step, a selection of meaningful indicators were further developed into a *Guidebook* for the organisation of the "Observatory of School Quality", a device to support schools in their self-monitoring of achievement and dropout. All schools involved in the programme "Education for All" were invited to implement the monitoring of their basic information system through the organisation of an "Observatory", consisting of a minimum of ten indicators to be selected from a set of 15 indicators organised in the following descriptive and evaluative dimensions:

- the family factors,
- the instructional factors,
- the educational context, and
- the results.

While the first model had been organised from a managerial perspective, this new version was prepared to respond to the teachers' view of the school. However, as

Table/Tableau 8.1.

Size of responding populations in two questionnaires on indicators
Nombre de répondants aux deux questionnaires sur les indicateurs

	Participation (7 schools)	Satisfaction (6 schools)
Teachers	83	93
Students	276	358
Other staff	56	31
Parents	543	305

important as the choice of the indicators was their organisational and conceptual framework made explicit in the evaluative dimensions.

In this new edition of the guidebook, some changes were introduced because some indicators had not seemed clear enough in their rationale and their method of computation. It seemed convenient to rephrase their descriptions or templates, add other arguments and revise the method.

The work developed so far responded to one of the purposes of the study: to select the performance indicators that best describe the school system, and to contribute to their improvement through the introduction of self-evaluation practices.

Accountability Indicators

The study of the sub-set of indicators that might be used as a common core in accountability processes and in schools' reporting to their communities, constituted the second step of this phase of the research. It seemed important to de-emphasise evaluation as a means of responding to the pressures for accountability coming both from the State and the society. In our opinion, that pressure in Portugal comes from the shortage of relevant and meaningful information about school performance which is accessible to large audiences. It therefore seemed useful to distinguish between the reporting activities that fulfil the demands of accountability, and the evaluation activities that support improvement, and may follow different methodologies, according to the purposes, the audiences and the available resources.

Having in mind the teachers' perceptions of "educational quality" and the results of our exploratory study, a model of indicators closer to the models of schools' productivity than to those of schools' effectiveness was developed. This covers seven "kernel" variables, corresponding to the description of three school dimensions: the results, the school context, including the students' family background and the school size, and the factors relating to teachers: their professional qualifications, experience and stability. The linguistic metaphor "kernel" was used to suggest the generative character of this model.

To gain a deeper insight into the effects of some contingent factors on school performance, two moderator variables were introduced: the type of school (C schools *versus* comprehensive C + S schools for the 2nd cycle; and comprehensive C + S schools *versus* secondary - S - schools for the 3rd cycle), and the geographical location of the school. The purpose was to observe the behaviour of the selected variables in the universe of Portuguese schools, and how they were mutually affected. For greater discrimination of the effects, an inter-quartile analysis was done. Furthermore, a cluster analysis procedure was designed to observe how the grouping of variables changed by the effects of the moderators.

3. Some Results

Some indicators in the list of the 50 initially selected could not be calculated. Either data were missing or they were not reliable. In some schools, entire areas of information disappeared at the turn of the school year: they were part of control mechanisms and having fulfilled their mission, they became useless, or just got lost. However, we could observe that some schools, namely the secondary ones, organised their information systems more easily than the others. Lower-level and smaller schools were neither autonomous in their own organisation nor systematic and complete in the gathering and analysis of data, though they possessed information covering aspects that were ignored by the others. Even in the more "exemplary schools" or in the most dynamic situations, data were used as discrete variables confined to their surface value.

In the first data collection (1990/91), our attention was concentrated on quantitative indicators, because their numeric nature made them easier to collect and treat. Teachers only had to mobilise existing figures. It was possible to compute 32 out of the 50 indicators we had proposed initially. Even these could not be treated in all the twelve schools. Nonetheless, they provided unknown information about many facts of school life that could be used to question school performance and effectiveness. There were many blank answers, but we tried to read and interpret what we had before us, being aware that most of the time more questions were asked than responded to.

During the second data collection (1991/92), some schools left the project. The remaining ones organised themselves differently for the task, and in general they constituted a task force for the school evaluation project, more or less dependent from the management team, according to local conditions. This time, qualitative indicators were also to be treated. In all schools, the questionnaires were addressed to a sample from each population, though the rigour in the sampling procedures had to be sub-ordinated to people's willingness to respond.

When the team returned to schools to discuss the data obtained so far, only some indicators were addressed, according to the teachers' interests or the school's needs. The order varied from school to school. However, the format of the discussion was very similar: the definition, the data, the justification or relevance of each indicator and the discussion of the findings. Generally, we moved from context to results; from results to functioning. "Resources" sometimes went together with "context"; at other times they went together with "functioning". However, our principal concern in these discussions

was to introduce "critical discourse" and reflective practices, not to cover all the indicators or inspect all the numbers produced. The "critical discourse" included the re-analysis of the enquiry tools, the discussion of the research-based arguments involved in the relevance of data, and the interpretation or meaning of the findings, as will be seen in the following three examples.

The Context Indicators

The context indicators were intended to account for the initial advantages or disadvantages of the schooling process. They could be organised in two sets: those related to the student's family background, or the school's external context, and those related to the school's organisational response, or to its internal context. They gathered information on the stability of the teaching force in the school, on the ratio of pupils to teachers or other adults, the school size and the provision of formal and informal education.

Among these indicators, "parents' level of education" generated greatest reflection, because of the span of the differences between schools, and because of the "real image of the country" which they conveyed, of which teachers seemed to be ignorant. "Parents' level of schooling" was defined as "the percentage of students whose father or mother had attained the 9th grade". Though we were interested in knowing the distribution of parents' levels of schooling, the indicator concerned just the 9th grade, which corresponds to the current level of basic compulsory schooling in Portugal. It was our belief that this indicator would provide useful information on the cultural level of our adult population and on the kind of efforts to be made to get a generation of students to reach that level. Teachers' attention was drawn to the fact that the socio-economic status of the families had been used as one of the variables to characterise the family background of the students. However, Walberg (1986) and Hattie (1987), in their studies on educational productivity, stress the importance of the characteristics of "the home environment" (cited by Fraser, 1989, p. 717), also called by others the "home curriculum", in child development and in the encouragement of achievement at school. Their attitudes, beliefs and expectations, as well as the intellectual and cultural stimuli they constitute, have a determining influence on students' cognitive and affective behaviour. In different ways, Fraser (1989) and Mortimore *et al.* (1989) underline the importance of "the academic stimulation of the home" (Fraser, 1989, p. 708) "the alterable curriculum of the home" (...) "twice as predictive of academic learning as is family SES" (Fraser, 1989, p. 717), or the "parental involvement in educational development within the home" (Mortimore *et al.*, 1989, p. 767). The high proportion of parents with only four years or less of compulsory schooling in the whole Portuguese population, and in the sample of twelve schools, is determinant of a common pattern of weakness in the cultural background of families in the public schools.

It is known among teachers that the weaker the families' cultural and academic standards, the greater their acceptance of their children's failure and the predisposition to accept the difficulty of school work or their children's lack of gifts. So it becomes the school's responsibility to support students' development and achievement also through planned education of the parents, explaining to them the intentions of the curriculum, the

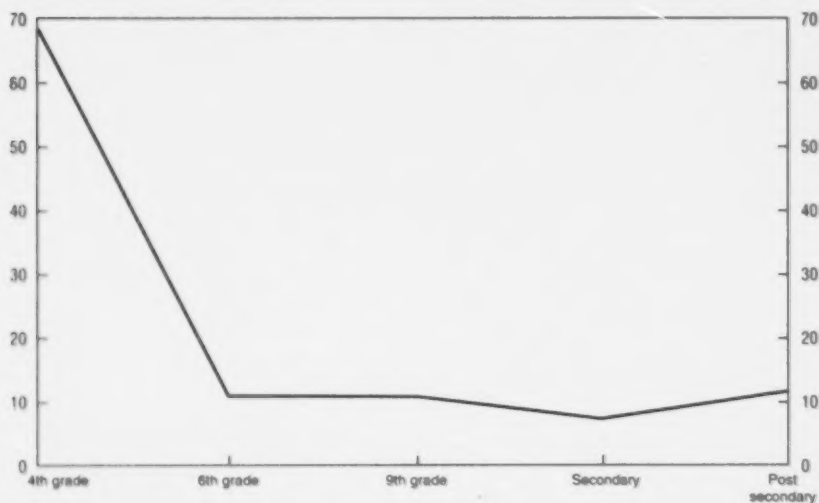
purposes of school activities and how they are to be developed, and the importance of their intervention (Figure 8.1).

It should be added that educational reforms go hand in hand with local cultural development. Therefore, reforms cannot be circumscribed by curricula and school life, but have to monitor the formal and informal relationships between schools and families and communities, or to use James Coleman's expression, have to control for the "social capital" of children and youngsters (cited in Stockard and Mayberry, 1992, p. 74).

The Results Indicators

During the discussion with schools on the selection of performance indicators, it was made clear that the results indicators refer to academic and non-academic results and cover a range of aspects related to the aims of education, even if for the moment we do not have reliable instrumentation or agreed criteria, and have difficulty in observing them. In the first year of the project, the academic results measured in schools were to be used, as well as other available measures related to the students' survival in the system. Though these describe aspects of the economy of the system, they can also be read as measures of

Figure 8.1. The profile of parents' level of education in the twelve schools in the sample
Niveau d'instruction des parents dans les douze établissements de l'échantillon



Source: Ministry of Education, Portugal.

the social effectiveness of the school. This is the case of school drop-out, school leavers' average age, and satisfaction.

The academic results were defined at two levels: the percentage of students passing each grade, every year; and the percentage of students passing and succeeding in all academic subjects. The former is the "rate of achievement"; the latter the "quality of achievement".

The traditional average measures hide important information on the quality of achievement and indirectly on the "accessibility of education". The study of a generation's academic survival in the education system contributed to a different view of the democratisation of schooling. Following a cohort of students enrolled for the first time in the first grade of school in 1982/83 till the end of the ninth grade of compulsory schooling in 1990/91, we verified that only 27 per cent of their generation successfully completed the nine years of basic education. The 14-year-olds, who in theory should have been leaving basic education, were spread all over the nine grades, accumulating a history of failure and dissatisfaction: 35 per cent were in the 8th or 7th grades; 14 per cent in the 6th and 5th grades; and 24 per cent had left school in the meantime, some of them prematurely entering the world of work, pushed by family interest and negative attitudes towards education. Data from individual schools were much more dramatic than the national picture, and teachers considered that this way of viewing the quality of the education system was rather brutal, conveying an image much less favourable than they were used to. But it made clear that some measures in use to describe the system may be illusory and do not contribute to changing schools: they have a weak diagnostic force and do not question existing educational policies and practices.

So the main focus of our analysis was "the quality of achievement", which seemed to be a good example of the illuminative potential of the measures of performance. Three modes of operationalising it were used:

- The relationship between "full" achievers and "flawed" achievers, meaning the percentage of the achieving students passing all subjects or failing one or more subjects. It was our argument that students passing despite failure in some subjects have more likelihood of being "failures" or "drop-outs" in the near future. Usually they are not considered an "at risk population" because they are counted in the group of achievers. But in a certain way they should be, if our concern is to guarantee the maximum educational level to larger strata of the population;
- The percentage of students failing in mathematics or the mother tongue, and cumulatively in both subjects. Teachers in general complain about the students' low mastery of reading and writing skills, which affects the whole learning process. Mathematics has the highest concentration of negative scores among all subjects. It was argued that difficulties in these two subjects can be attributed to students but also to the organisation of schools and the teaching process, including the design of curricula. However, if attributable to the students, probably the failure would be cumulative;
- The distribution of the levels of achievement in Portuguese, mathematics and foreign languages. We suspected that these three subjects illustrated three different attitudes to work and to learning, which are mirrored in the teachers' classroom

work, in the materials they use, in how they deal with homework, and in the way in which teachers mark their students' work.

Data obtained so far confirm the importance of this indicator. In 1990/91, in the second cycle (9-11 years of age), there were 82.6 per cent of achievers in the whole population of the six schools in the project which offered this level of studies. However, of this percentage, only 52 per cent were "full achievers", that is, passed in all subjects. It means that 48 per cent of the 2nd cycle achievers may constitute, in spite of apparent success, a population at risk of further failure and drop-out, because of their greater difficulties in learning and natural demotivation. Research says that prior knowledge is a correlate of achievement (Fraser, 1989), and that weaknesses in previous learning are associated with other causes of school drop-out and early labour market entry (Ferrão, 1992, p. 77), with the well-known social consequences. Therefore, it is not surprising that in the following cycle, between the 7th and the end of the 9th grade, 33 per cent of the youngsters in that cycle have abandoned school in Portugal (1989/90, 1991/92).

The analysis of the 77.3 per cent of the whole population of the six schools in our sample that achieved in the 3rd cycle, shows that only 41 per cent of them were full achievers, which forecasts a reduced school-life survival at the secondary level for a significant proportion of our students.

In 1991/92, achievement rates increased as a percentage, but the relationship between the achievement rates and the quality of achievement remained the same.

The Factors Related to Teachers

Teachers are unanimously considered to be key factors in the quality of schooling. However, most of the indicators concerned with teachers are designed from the perspective of human resources analysis, the economy of education and managerial concerns. Trying to derive from the initial system of indicators a subsidiary conceptual model, we approached the factors concerned with teachers from three angles: as human resources, as elements of the environment, and as process or functioning agents of school life. Data had been collected in all schools to describe the range of difference in staff quality, covering aspects such as "teachers' qualifications", "experience", "seniority", and "in-service training". These characteristics may describe some attributes of human resources in education. From this perspective, teachers are viewed as "inputs" of the system, and indicate the amount of effort expressed in terms of energy and cost which are or would have to be made in order to have a qualified staff in a given school.

Elements such as the ratios of students to teachers and of students to adults, and the stability of the teaching force in each school, traditionally presented as economic indicators of the education system, were viewed rather as features of the schools' internal context. In other words, they were seen to refer to the internal arrangements that may constitute specific and varying conditions of school functioning, an infrastructure of the quality of schooling. From this perspective, they are fundamental factors of the learning environment.

The way in which teachers and others involve, react to and "consume" school life was approached through school functioning variables, namely "satisfaction" and "par-

ticipation". The other indicators of functioning were analysed from the point of view of the description of the place of work: the level of autonomy, the management style, and the communication and information modes at school level.

Two aspects of the factors related to teachers will be illustrated: teachers as "human resources", and "participation" as a feature of their quality as agents and consumers of school life. (Information on students' and parents' participation will be included to compare three different patterns of participation.)

Teachers as Human Resources

Three indicators were selected: "teachers' professional qualifications", or the percentage of teachers having a professional degree in teaching; "teachers' experience", or the proportion of teachers having ten years of teaching experience or more; and "stability of the teaching force in schools", or the percentage of teachers working in the school for ten years or more.

Figures showed that human resources vary from school to school. Moreover, they can change drastically from one year to the next: the stability of the teaching force, the demography of students, the curriculum offered, and the strategies of organisation are at the origin of some of the differences observed. An experienced and qualified teaching force was seen in some schools, while in others a large proportion of teachers were young, beginning teachers. Taken individually, teachers' qualifications are not a correlate of school effectiveness, just as no other resource is. However, in combination with other characteristics of schools, they constitute one of the "enabling conditions" of students' achievement and as such they can be taken as one of the measures of "knowledge accessibility" (Oakes, 1989, p.193) in the analysis of a school's context.

A detailed analysis of the teachers' specific qualifications provided a clearer picture of the differences. In all schools, a higher proportion of the teachers of Portuguese had higher degrees than the respective teachers of mathematics. In this case, the proportion of unqualified teachers was alarming, especially as they were concentrated in specific schools, all located in peripheral or small urban areas.

Participation in School Life by Different Groups

The study of "participation" was intended to explain how and how frequently the teachers and the other different populations acting at school level, participated in school life, namely in the areas that norms and tradition have established as privileged, common or specific fields of participation.

Participation has been considered a feature of effective schools and one of the components of school improvement, since it is "systematically related to the quality of the school as a social system" (Bollan and Hopkins, 1988, p. 19). White (1992) refers to the impact of participation on teachers' working conditions and specifies as its "major benefits a) improved teacher morale, b) better-informed teachers, c) improved teacher communication within and across schools, d) improved student motivation, and e) increased incentives that serve to attract and retain quality teachers" (White, 1992, p. 71).

Participation in decision-making contributes to increased self-esteem, and the sense of school ownership. Being a precise mode of social behaviour, it fosters the sharing of expertise, "interactive professionalism" (Fullan, 1991, p. 4) and is associated with the generation "of a culture of inter-professional development" (Simons, 1988, p. 15).

Participation has also been referred to as a condition for the strengthening of the different actors' motivation and cohesion. While motivation implies the individual willingness to be involved in specific situations or acts, cohesion is an attribute of social behaviour characterised by processes of decision-making or implementation of any project. Cohesion goes together with collegiality. From the perspective of collegiality and cohesion, participation can be thought of as an expression of institutional solidarity and a prerequisite to the learning of citizenship.

Students' participation has been referred to in some literature as a means of broadening their school experience, having "significant consequences on later cognitive, social and personality development" and on later survival in academic life (Schoggen and Schoggen, 1988, p. 293).

The arguments for parents' participation are different. In the history of their involvement with school life, educational reasons have dominated. Their responsibilities as the first educators and their role as supporters of the action of the school reflect the importance of mutual interaction for the improvement of students' performance: their integration, their emotional and affective development and their achievement. Recently, economic and political reasons have been advanced to justify the parents' role as partners in education and their involvement in the orientation of local educational policy, in decision-making and in the control of the use of resources and of results.

"Participation" was conceptualised in our study from three points of view, or as three modes: participation in the "presentation of work or action proposals" (presentation); participation in the "discussion of action proposals and suggestions" (discussion), and participation in "decision-taking" (decision) on the respective proposals. Each mode corresponds to a qualitative difference in the actors' engagement or in the conceptual level of their participation in school life.

Four sets of similar questionnaires were developed addressing four different populations: the teachers, the students, the other school staff and the parents. In total, 958 subjects from seven schools were involved.

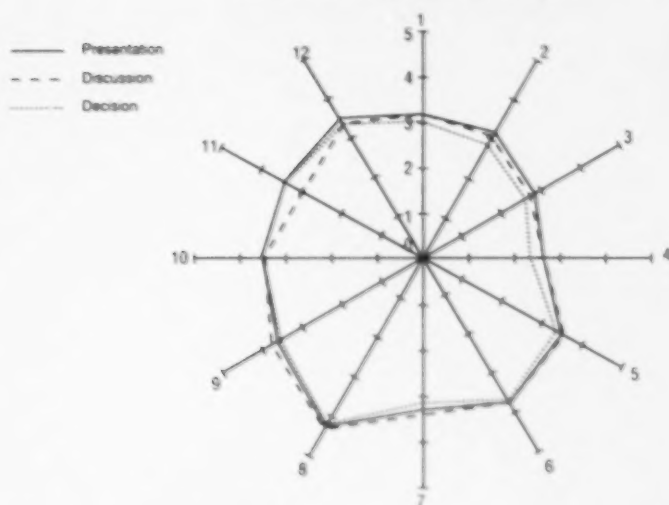
Participation by Teachers

Teachers were questioned on two basic areas or domains: "the institutional life domain" and "the teaching process domain". The former addressed the issues of school rules, the setting of school goals, the discipline problems and the budget forecast and investment. The latter covered items related to the curricula, the processes of assessing students, and the non-formal (or non-classroom) activities.

The analysis of the main tendencies observed in the data from the 83 observations, indicated (see Figure 8.2) that:

- Teachers participate at a moderate rate in school life (3.4 on a frequency scale of 5);

Figure 8.2. The teachers' modes of participation
Modes de participation des enseignants



Source: Ministry of Education, Portugal.

- They perceive three distinct "objects" of participation: the organisation of institutional life; their professional activity, and the assessment of students;
- In general, the organisation of institutional life does not involve teachers very strongly. "Discipline" is the issue in this domain that conveys higher levels of participation;
- The items related to the "teaching process", or to teachers' professional activity, constitute the domain with the highest rates of participation, whatever the mode considered;
- "Textbook selection" is the topic that scores highest in participation, on any mode;
- There is greater variation in the rates of teachers' participation in what concerns the "objects" than in the "modes" of participation. However, they participate more frequently in "presenting" and "discussing", than in deciding on school matters. This may mean that the teachers' real power in schools is less than their capacity of intervention.

Participation by Students

The questionnaire addressed to students was a short one covering two domains: the "institutional domain" including the "making of school rules" and "discipline issues", and the learning process or the "educational domain", covering the "setting of learning targets", the "organisation of cultural activities" and the "organisation of study visits".

According to the data (involving 276 subjects), students' participation in school life is low, almost non-existent (2 on a frequency scale of 5). The low homogeneity in their ratings suggests that their participatory experiences probably vary from school to school. In spite of the low rate of participation, we can distinguish a relatively greater involvement in the planning and organisation of school experiences than in the formal, norm-based, functioning of school. Students are thus sometimes or often involved in the process of "planning and organising study visits", either through making suggestions, discussing or taking decisions on them; in the "organisation of cultural activities"; or in the "setting of learning targets". On the other hand, they "seldom" or "never" participate in the "making of school rules" or in the discussion of "discipline issues". As we can see in the diagram (Figure 8.3), students' participation varies with the "objects of participation" and with the respective modes.

If we compare the students' and the teachers' responses we can observe two main differences: teachers' participation follows the same pattern in the three modes considered, the differences between the areas of concern being a matter of degree. This means that participation by teachers may be conditioned by the "objects", not by the modes of participation: teachers may participate at a low rate, but when they do, they follow the whole process, from presentation to decision-taking. This reveals the existence of a participatory policy, which may be implemented more or less thoroughly in each place, according to the people and their values.

In the students' case it varies: they may be involved in the various phases of the decision process, but there seem to be two distinct attitudes concerning students' participation. On the one hand, students are accepted as initiative takers and they may even be invited to give school administrators their opinions on various topics. However, they are essentially viewed as clients or consumers of the educational services, and when they are consulted, it is to be better served, not to assume responsibilities. Thus, they may present opinions and suggestions, and discuss others' proposals, but they do not decide. On the other hand, from 3rd cycle schools onwards, according to law, they are "partners" in specific areas of school organisation, and theoretically their participation takes the form of co-responsibility. They are supposed to discuss and decide. This attitude corresponds to a truly participatory policy. One example is the students' participation in school pedagogical councils and disciplinary councils, as has been enshrined in law. Participation is a right and a strategy for development, not an award for maturity. This distinction corresponds to two different views on the role of participation in school life and in school improvement.

The scores from the questionnaires seem to indicate that the students' experience is coloured by some contradiction in the understanding of the modes of their participation: they are not "deciders", and even the room for them to "discuss" problems is very

Figure 8.3. The students' modes of participation
Modes de participation des élèves

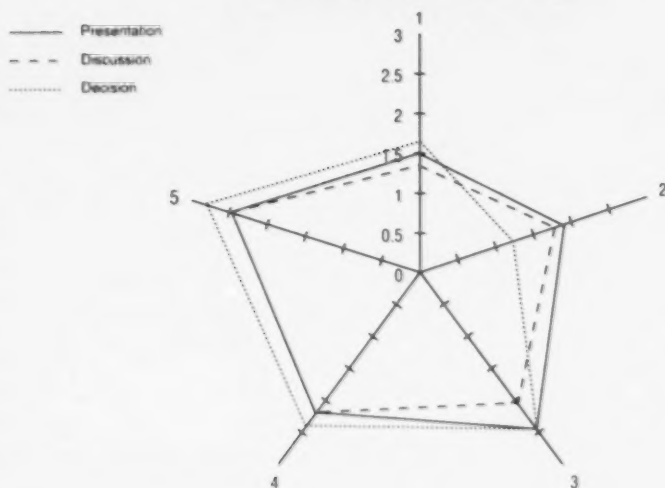
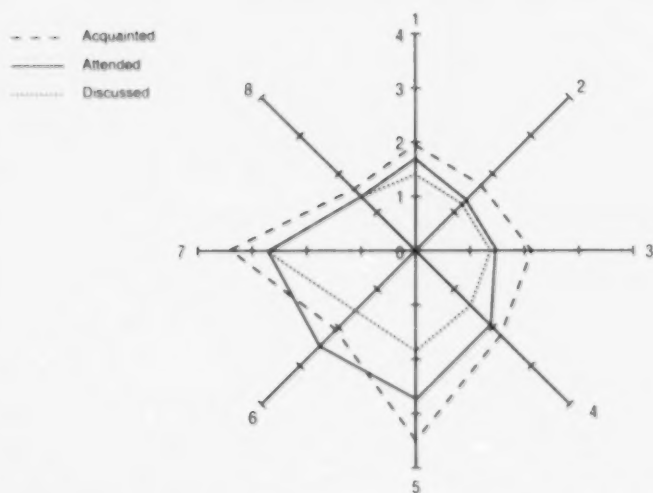


Figure 8.4. The parents' modes of participation
Modes de participation des parents



Source : Ministry of Education, Portugal.

limited in many places. However, as can be seen in the diagram (Figure 8.3), they "decide" more than they discuss.

Participation by Parents

Parents' involvement with schools is related to their role in the educational process and may follow distinctive participation models, each of them more related to the processes of information and communication than to decision-making. Three modes or levels of participation were therefore considered: at a first level, parents are viewed as the addressees of uni-directional information, usually the written information which schools produce, and which only require parents to be informed of, or acquainted with specific issues, without requiring them to give schools their reactions. At the other two levels of participation, parents are perceived as "helpers", persons to be involved to provide support or to legitimate the actions of others, and as "partners" (Davies, 1986, p. 120), persons to share responsibility in the taking of decisions.

As "helpers", they are expected to be involved and to react actively to schools' information and requests. Parents are requested to give schools opinions and suggestions, to participate actively by attending meetings, helping in the learning processes, discussing and preparing the decision-making with teachers or within their organisational groups. The information is bi-directional, written or oral. As "partners", parents are expected to be members of the educational community, to take initiatives, to discuss and prepare solutions, to influence the direction taken by the school, to be opinion-makers.

Three modes of participation were conceived: to be informed or acquainted with school issues; to attend meetings without involvement; and to discuss the issues under analysis. Data from 543 subjects revealed that parents' levels of information on school life is low except when they are personally concerned, or when the topics concern their own children, and they seldom participate in meetings to discuss school problems, unless they are personally affected. As we can observe in Figure 8.4, parents' participation varies with the "objects" or with the issues to be dealt with, but it also varies with the modes of participation. The "participation polygon" illustrates how weak their role as partners is, except when it concerns their own child's achievement. The parents' participation as partners would presuppose their involvement in the discussion of and in the choices about institutional life at a level similar to their involvement in the other modes.

The existence of a positive participatory culture would be pictured by overlapping polygons, as we have seen in the images of teachers, whose participation is relatively low; the strategies to increase it should concern "intensification" (Fullan, 1992, p. 1). On the other hand, the strategies to increase parents' participation will imply a cultural change at school level to be sustained by "restructuring" initiatives, which, *per se*, only make participation functionally possible, but do not diminish the conflictual interests of parents and teachers, nor their self-centredness in their relationships with school life. Only a culture of interaction will promote different patterns of participation.

Other examples could be given of how the different indicators were approached, how the information was mobilised or how it related to school improvement. Though the process changed from school to school, depending on the local concerns with achievement and quality of school life, and on the local capacity to view education as policy-

making, the intention was to move from the figures and the percentages game to the search for the meanings that could be derived from numbers and to relate them to educational action at the individual and institutional levels.

4. From Improvement to Accountability

The discussion of the performance indicators was an opportunity for schools to exchange information and talk to each other in an organised and systematic way, centred on the data and the procedures to collect them and compute the indicators. These interactions made everyone ask how far each of the schools scored above or below the national average on the different indicators, and what would be the range of variation between schools. Usually, the statistical information used to assess the education system's performance takes individuals as its referents, the main units of measurement being the students and the teachers. However, these measures seemed inaccurate to describe and monitor the performance of the organisational units of the system. Moreover, the traditional measures do not take into account the differences in the contexts of schools, in functioning and in results, and they are essential for policy-making, for planning and for judgement.

The need was felt to use multiple referents so that the complexity of schools might be minimally captured, and the evaluation of the quality of schools on the basis of single criteria be avoided. The search for "sensitive" referents raised two kinds of problem: firstly, the measures to be considered "sensitive"; secondly, the criteria to be used in the selection of evaluative variables.

The review of the work developed so far made everyone aware of the differences between indicators in terms of relevance, utility and meaningfulness. Many of them had encountered problems in the collection of data, in that the schools did not have the equipment, time, and expertise to do the job. Only 32 indicators had been computed, 14 of which seemed to be statistically meaningful: the rate of achievement and the quality of achievement correlated highly with the family background factors; and the functioning variables such as the time on task, the safe atmosphere, and the dynamics of schools correlated with human resources, *i.e.* professional qualifications, teachers' experience and the stability of the teaching force, teachers' involvement with training, and organisational complexity. These indicators were among the ones that teachers rated as "the most important".

The work developed so far had an exploratory character and essentially constituted a way of approaching self-monitoring by schools as a means of supporting their improvement. However, it seemed important to establish the difference between the organisation of the schools' information system around a model of performance indicators, as an approach to systematic self-regulation and to the activities of self-evaluation, and the organisation of accountability processes. On the one hand, schools need elements for making judgements on aspects related to their policies and goals, *i.e.* on their educational project. Those elements of information may be different over time, according to the stage of the project's implementation. On the other hand, they need stable elements of information to be used in reporting on their performance, which can give an account of their

progress. These elements of information need to be framed by a synthetic model so that the selection is coherent and the community at large may process the information, and the message.

Development of a Common Core of Indicators

It was thought that this model might constitute a common core in schools' information systems and accountability processes: a common core, because all schools would be involved in that reporting, and would use the same language to introduce logic and consistency into the comparisons to be made.

Seven factors were selected for this purpose: achievement rate, quality of achievement, teachers' qualifications, teachers' experience, stability of teaching staff, and parents' level of education. Teachers also considered that school size was an important factor affecting students' performance and teachers' working conditions. Research has lent support to this argument, namely the negative effects of size on performance in high SES contexts (Fowler and Walberg, 1991; Fetler, 1989; Friedkin and Necoechea, 1988). For practical reasons, no qualitative factor was added, since there was no available information in the national data banks. The system of indicators therefore became closer to educational productivity models than to descriptors of school effectiveness, but followed a logic of school improvement (Lagerweij and Voogt, 1990).

Testing of the Common Core of Indicators

To test the strength of this model in describing the distribution of school performance by means of the seven selected factors, an inter-quartile analysis was conducted, so that a more discriminating picture of the different levels of the education system could be made, in order to allow for more effective strategic planning in educational action. The purpose was to illustrate within what parameters the "worst" schools, those falling in the first quartile, scored, and how far they were from the "top" ones, the 25 per cent with the highest scores.

There was the feeling that all these factors could be affected by contextual or contingent variables such as the type of school for each level of schooling, and the geographical zone where the school is located: urban areas, and the different categories of rural zones. The second cycle may be delivered at C schools (intended for the 2nd cycle only) or at C + S schools, which correspond to comprehensive schools. The 3rd cycle may be delivered at C + S or at secondary schools (S). With regard to geographical zones, we followed the classification of municipalities published by the Ministry of Internal Affairs and the Administration of the Territory, which controls for two variables, population size and taxes collected. They give five categories of municipalities: the urban, and the 1st order rural, the 2nd, the 3rd and the 4th order rural. Because of the size of the sample of schools in each category, we put together the 2nd and 3rd order rural zones. Type and zone variables are related to differences in history, in culture, in attitudes and in demography.

Four models of analysis were thus designed for the observation of the performance of the school system at the national level: the universe (model 1), the type of school

(model 2), the geographical zone (model 3) and the type and zone cumulatively (model 4). For a greater discrimination of data, an inter-quartile analysis of performance data for each model was conducted. Both moderator variables were shown to affect schools in the different quartile intervals in different ways. Model 4 proved to be the most relevant, since it showed that "type" and "zone" have a contingent moderator role: they affect performance, but only under certain conditions.

For example, in the second cycle, level C schools score higher on all indicators than C + S, except in "size", and in all quartile intervals.

In the third cycle, C + S schools score higher in students' results than secondary schools (S), although the latter have higher percentages of qualified, experienced and long-serving ("stable") teachers, and a higher percentage of parents educated to the 9th grade or above. This reminds us of the tradition of selectivity of former *lycea* (current secondary schools), whose tradition of quality was based mostly on the qualifications of their human resources and on high achievers.

If these indicators are broken down by "type" and "geographical zone", the 3rd cycle schools located in the 1st and 2nd order rural areas score higher in the two students' achievement indicators than the urban and the 4th order rural schools, in all quartile intervals. This suggests that for various reasons, schools located in these geographical areas need differentiated support, as they are affected by different social factors. More qualified teachers and parents tend to settle in similar places, where life may have some quality, but other factors affect school functioning, making the achievement more problematic.

To illustrate the differences in the distribution of the performance levels by quartile intervals, in the seven selected indicators, the scores of the 2nd cycle schools are displayed in Table 8.2.

Table 8.2 shows, from the perspective of monitoring school improvement, that it is important to have an idea of where each school falls in the distribution of performance scores. The quartile interval measures make clearer the efforts necessary to move from one interval to the other than the traditional national averages do. They seem particularly helpful in appraising performance, and provide a more accurate image of each school.

Table 8.3 is an example of the fact that in school improvement strategies and in accountability it is important to consider certain contextual variables that affect performance. These are beyond the control of schools and highlight how schools operate in different conditions and with different results.

The importance of considering the interaction of the different factors in different contexts is confirmed by a cluster analysis, which highlights how the seven indicators can be differently grouped according to the contextual variables, and depending on the level of performance, or the quartile interval considered.

Thus, for example, in the 2nd cycle schools, the "quality of achievement" clusters together with the "rate of achievement" in C + S schools, and both variables constitute two independent clusters in C schools. The same can be observed in the different geographical zones, except in the most deprived areas. Schools in the 3rd and 4th order rural zones cluster the "quality of achievement" with the "parents' level of education"

Table/Tableau 8.2.

**Quartile distribution of average performance in the universe
of second cycle schools, 1990/91**

Per cent, N = 574 schools

***Distribution par quartile des résultats moyens pour l'ensemble des établissements secondaires
du deuxième cycle, 1990/91***

Quartile interval	Rate of achievement	Quality of achievement	Teachers' qualifications	Teachers' experience	Stability of teaching force	Parents' education	School size
0	0.78	0.47	0.46	0.35	0.29	0.05	401
1	0.84	0.55	0.71	0.55	0.58	0.12	624
2	0.88	0.61	0.83	0.70	0.73	0.21	888
3	0.93	0.72	0.93	0.85	0.88	0.40	2 864
Average	0.86	0.59	0.73	0.61	0.62	0.20	

Definitions:

Rate of achievement: Percentage of the enrolled students achieving in the cycle.

Quality of achievement: Percentage of achievers succeeding in all subjects.

Teachers' qualifications: Percentage of teachers having a professional degree in teaching.

Teachers' experience: Percentage of teachers in the career for ten or more years.

Stability of the teaching force: Percentage of teachers working in the school for three or more years.

Parents' level of education: Percentage of students having at least one parent educated to 9th grade or above.

School size: Maximum number of enrolled students in the schools at each interval.

Note: In the 25 per cent of the second cycle schools with the lowest performance scores (interval 0), only 78 per cent of the students passed. In the 25 per cent of the "top ones" (interval 3), that value increased to 93 per cent of the students.

Table/Tableau 8.3.

**Quartile distribution of performance in "Quality of Achievement"
in the second cycle schools,
controlling for "Type of School" and "Geographical Zone", 1990/91**

***Distribution par quartile de la performance concernant « la qualité des résultats scolaires »
dans les établissements d'enseignement secondaire du deuxième cycle,
compte tenu des « types d'établissement »
et des « zones géographiques », 1990/91***

Quartile intervals	C/urban	C/rural 1st	C/rural 2nd + 3rd	C/rural 4th	C + S/urban	C + S/rural 1st	C + S/rural 2nd + 3rd	C + S/rural 4th
0	0.49	0.51	0.51	0.34	0.46	0.49	0.45	0.44
1	0.56	0.58	0.59	0.57	0.52	0.57	0.54	0.55
2	0.61	0.63	0.64	0.62	0.58	0.62	0.61	0.62
3	0.70	0.72	0.76	0.72	0.66	0.78	0.74	0.71
Average	0.59	0.61	0.63	0.56	0.56	0.62	0.59	0.52

Note: In the 25 per cent of the second cycle urban schools (C/urban) with the lowest scores in "Quality of Achievement" (interval 0), only 49 per cent of students succeeded in all subjects. In the 25 per cent of the best C/urban schools (interval 3), 70 per cent of the achievers succeeded in all subjects.

on scores in the first quartile interval, or cluster the "quality of achievement" with the "school size" on scores in the fourth quartile. In both cases, the "rate of achievement" constitutes an independent cluster, geometrically distant from the others.

However, when the performance indicators are broken down by "type" and "zone" cumulatively, different patterns of clustering are obtained, stressing the importance of these contextual variables in the discrimination of school quality. In monitoring and in appraisal of educational performance, if they cannot be used as predictors of the quality of functioning, they are to be considered as factors of differentiation.

5. Conclusion

This chapter has reported a study developed in Portugal on the use of performance indicators as an approach to school improvement, and to accountability, involving a group of twelve schools from the different schooling levels and a research team in a central department in the Ministry of Education: the Department of Programming and Financial Management.

Having been implemented in parallel to the INES project, it benefited from the work developed in Network C. Some of the purposes coincide, namely those referring to:

- the consolidation and the refinement of data and data-gathering processes in schools;
- the enrichment of conceptual work on the development, presentation and interpretation of indicators;
- the comparison between schools;
- the analysis of the ways in which systems function.

Even the format of some of the tools used in this national study correspond to a simplified version of some of those produced by Network C.

The organisation of schools' information systems around a model of performance indicators referred to as *relevant* in the literature of school effectiveness, as *important* in administrators' discourse, and as *useful* in teachers' and school managers' opinion, constituted the first step in the work to be developed. The computing of school indicators, and the analysis of their informative capacity to support school improvement, responded to the major purposes of the study.

The development of a critical discourse around the indicators computed aimed at supporting the access to research-based thinking, and the commitment to rigour and rationality. It also constituted the approach to schools' self-evaluation: "The collection, analysis, interpretation and use of comprehensive information can play a major role in evaluation for local school improvement" (Sirotnik, 1987, p. 88).

Some examples have been given of how the indicators were approached in the dialogues between the team of researchers and the teachers in school meetings.

The making of demands on accountable schools has been the other face of the granting of autonomy. Usually, the central authorities define the contents, the pace and

the format of accounting procedures, and this reporting is perceived by public opinion as an evaluation procedure of the education system.

The research developed so far proposes clarification between the regular activities of reporting to the community and those concerned with the appraisal of performance. It suggests a bottom-up approach to accountability, generating from the initial model of performance indicators a synthetic one that controls for essential dimensions and characteristics of schools.

These seven indicators may constitute the referents for judgement of the performance of the education system and the quality of each individual school. It is up to each school to add to this core the indicators that best respond to the features of its own plan of development. We only propose some examples of such a selection.

References

- ARDOINO, A. (1980), *Éducation et relations: introduction à une analyse plurielle des situations éducatives*, Gauthier-Villars/UNESCO, Paris.
- ARGYRIS, C. and SCHÖN, D. (1978), *Organisational Learning: a Theory of Action Perspective*, Addison Wesley Publishing Company, Reading, MA.
- AUBÉGNY, J. (1989), *L'évaluation des organisations éducatives*, UNMFREO, Éditions Universitaires, Paris.
- BERNARD, M. (1991), *The Evaluation of Schooling Institutions*, Paper presented at a conference at the Universidade Nova de Lisboa, Lisbon.
- BLANK, R. (1993), "Developing a system of education indicators: selecting, implementing, and reporting indicators", *Education Evaluation and Policy Analysis*, Vol. 15 (1), pp. 65-80.
- BOLLAN, R. and HOPKINS, D. (1988), *La pratique de l'auto-analyse de l'établissement scolaire*, OECD/JSIP, Economica, Paris.
- CARNEIRO, R. (1980), "Planification de la réforme de l'éducation", *Les réformes de l'éducation: expériences et perspectives*, UNESCO, Paris.
- COMFORT, L. (1982), *Education Policy and Evaluation: a Context for Change*, Pergamon Press, Oxford.
- COSTA, L. (1981), "Função social da educação", *O Sistema Educativo*, Fundação Calouste Gulbenkian, Lisbon.
- CRANDAL, D., EISEMAN, J., and LOUIS, K.S. (1986), "Strategic planning issues that bear on the success of school improvement efforts", *Educational Administration Quarterly*, Vol. 22 (3), pp. 21-53.
- da CUNHA, P. (1991), *Increasing Democracy in Schools as an Imperative Factor of Educational Reform*, Paper presented at the 2nd European Conference on School Democratisation, Final report, GEP.
- DAVIES, E. (1986), "The changing roles of parents, governors and teachers", *Educational Management and Administration*, Vol. 14 (2), pp. 120-124.
- FERRÃO, J. (1992), *Caracterização regional dos factores de abandono escolar nos 2º e 3º ciclos do ensino básico*, PEPT 2000, Ministry of Education, Lisbon.
- FETLER, M. (1989), "School dropout rates, academic performance, size and poverty: correlates of educational reform", *Education Evaluation and Policy Analysis*, Vol. 11 (2), pp. 109-116.
- FIDLER, B. and FUGL, B. (1991), "Managing in the 1990's", *Educational Management and Administration*, Vol. 19, No. 2, pp. 66-70.
- FOWLER, W. and WALBERG, H.J. (1991), "School size, characteristics and outcomes", *Education Evaluation and Policy Analysis*, Vol. 13 (2), pp. 189-202.

- FRASER, B. (1989), "Research synthesis on school and instructional effectiveness", *International Journal of Educational Research*, Vol. 13 (7), pp. 707-718.
- FRIEDKIN, N. and NECOCHEA (1988), "School system size and performance: a contingency perspective", *Education Evaluation and Policy Analysis*, Vol. 10 (3), pp. 237-249.
- FULLAN, M. (1991), "Education reform and teacher professionalism", Paper presented in the Australian Council for Educational Administration, Conference in Southfort, *Network News International*, Vol. 11 (4).
- FULLAN, M. (1992), *Successful School Improvement, the Implementation Perspective and Beyond*, Open University Press, Milton Keynes.
- HOROWITZ, C. (1990), "Total quality management: an approach for education?", *Education Management and Administration*, Vol. 18 (2), pp. 55-58.
- JONES, K. and MITCHELL, J. (1990), "Communicating evaluation findings: the use of a chart essay", *Education Evaluation and Policy Analysis*, Vol. 12 (4), pp. 449-462.
- LAGERWEIJ, N. and VOOGT, J. (1990), "Policy making at the school level: some issues for the 90's", *School Effectiveness and School Improvement*, Vol. 1, No. 2, pp. 98-120.
- MAYSTON, D. et al. (1991), "Educational performance assessment - a new framework of analysis", *Policy and Politics*, Vol. 19 (2), pp. 99-108.
- MCLEOD, G. (1988), "Teacher self-evaluation: an analysis of criteria, indicators and processes used by teachers in judging their success", *International Journal of Educational Research*, Vol. 12 (4), pp. 395-407.
- MFA (1974), *Programa do Movimento das Forças Armadas*, MFA, Lisbon.
- Ministério da Educação Nacional (1973), *A reforma do sistema educativo*, Ministry of Education, Lisbon.
- MORTIMORE, P. et al. (1989), "A study of effective junior schools", *International Journal of Educational Research*, Vol. 13 (7), pp. 753-768.
- MURNANE, R. (1987), "Improving education indicators and economic indicators: the same problems?", *Education Evaluation and Policy Analysis*, Vol. 9 (2), pp. 101-116.
- OAKES, J. (1989), "What educational indicators? The case for assessing the school context", *Education Evaluation and Policy Analysis*, Vol. 11 (2), pp. 181-199.
- OECD (1984), *Reviews of National Policies for Education: Portugal*, Paris.
- RICHARDS, C. (1988), "A typology of educational monitoring systems", *Educational Evaluation and Policy Analysis*, Vol. 10 (2), pp. 106-116.
- SCHEERENS, J. (1990), "School effectiveness research and the development of process indicators of school functioning", *School Effectiveness and School Improvement*, Vol. 1 (1), pp. 61-80.
- SCHOGGEN, P. and SCHOGGEN, M. (1988), "Student voluntary participation and high school size", *Journal of Educational Research*, Vol. 81 (5), pp. 283-293.
- SIMONS, H. (1988), "Evaluation and the reform of schools", Paper presented at the conference on the Evaluation of Educational Programmes: Methods, Uses and Benefits, North Berwick, Scotland.
- SIROTNIK, K. (1987), "The information side of evaluation for local school improvement", *International Journal of Educational Research*, Vol. 11 (1), pp. 77-89.
- SPARKES, A. (1991), "The culture of teaching, critical reflection and change: possibilities and changes", *Educational Management and Administration*, Vol. 19 (1), pp. 4-19.

STOCKARD, J. and MAYBERRY, M. (1992), *Effective Educational Environments*, Corwin Press, Inc.

WHITE, P. (1992), "Teacher empowerment under 'ideal' school site autonomy", *Education Evaluation and Policy Analysis*, Vol. 14 (1), pp. 69-82.

Chapter/Chapitre 9

**Perspectives on Decision-Making in Education Systems and the
Development of the Network C Locus of Decision-Making Survey**
*Considérations sur la prise de décision dans les systèmes
d'enseignement et sur l'évolution de l'enquête menée par le Réseau C
sur les niveaux où se prennent les décisions*

by

Darrel Drury and Laura Hersh Salganik
Pelavin Research Institute, Washington, DC, United States

and

Marilyn M. McMillen
National Center for Education Statistics, US Department of Education,
Washington, DC, United States

In 1990, Network C of the OECD INES project identified school-level decision-making as a topic of interest in Member countries. An instrument to collect information from experts – the Survey on Locus and Modes of Decision-Making – was piloted and subsequently revised, and the resulting indicators focusing on school-level decision-making were included in *Education at a Glance* (OECD, 1992 and 1993a). The purpose of this chapter is to introduce ideas for building on the accomplishments of the initial survey. The first section discusses two areas of organisational theory, selected to provide a conceptual framework for schools as organisations and for decision-making itself. The second section suggests additional interpretations of the current results and areas for expansion of the survey. The illustrative examples are taken primarily from the United States.

Institutional theory notes that the regulation of categories which are highly associated with institutionalised norms of behaviour and values allows education systems to shelter their technical activity from organisational controls. According to political systems theory, however, organisational participants should be viewed as political actors with their own needs, objectives, and strategies to achieve those objectives. A synthesis of the theories suggests that it is the differential ability of divergent interests to shape,

interpret, and manipulate the institutional environment that, in large part, determines which organisational forms are adopted and which are not. Suggestions for expanding the survey include the incorporation of: informal influence processes; extra-organisational actors such as the public, parents, and unions; the breadth of decision-making frameworks; the extent to which advice given in consultation is followed; decision-making processes within as well as across levels; and the degree of implementation of decisions. It is also proposed that decisions regarding evaluation of teachers, schools, and students should be added to the survey.

*
* *

Note de synthèse

En 1990, le Réseau C du projet INES de l'OCDE a constaté que la prise de décision au niveau de l'école était un sujet qui intéressait les pays Membres. Un questionnaire d'enquête destiné à recueillir des renseignements auprès des experts en la matière – l'Enquête sur les niveaux et modes de décision dans les systèmes éducatifs – a été appliqué et révisé par la suite, et les indicateurs qui en résultent, et qui portent sur la prise de décision au niveau de l'école, ont été inclus dans Regards sur l'éducation (OCDE, 1992 et 1993a). Ce chapitre a pour objet de proposer des moyens d'aller plus loin à partir des résultats de la première enquête. Dans la première section, on étudie deux domaines de la théorie de l'organisation, qui forment le cadre théorique à l'intérieur duquel on peut analyser les écoles perçues en tant qu'organisations ainsi que la prise de décision proprement dite. Dans la deuxième section, on propose d'autres interprétations des résultats actuels ainsi que des domaines auxquels pourrait s'étendre l'enquête. Les illustrations proviennent principalement des Etats-Unis car c'est le système éducatif américain que nous connaissons le mieux.

Dans la théorie institutionnelle qui est présentée en premier lieu, on note que la division en catégories (tels que les classes, les matières et disciplines, les cours, les enseignants et les élèves) qui sont étroitement liées aux normes institutionnelles de comportement et de valeurs, permet aux systèmes éducatifs de mettre leur activité technique (l'enseignement dispensé aux élèves) à l'abri des contrôles administratifs. Les niveaux supérieurs du système s'organisent en partant du principe que les niveaux inférieurs fonctionnent bien du moment qu'ils se conforment à ces normes et catégories institutionnelles. En conséquence, il y a relativement peu de vérification directe de l'enseignement dispensé et d'inspection directe des résultats. En outre, le comportement des acteurs de l'organisation est influencé autant par les attentes institutionnelles que par les politiques élaborées au moyen des processus de prise de décision.

Pour comprendre l'évolution et le fonctionnement des structures de décision dans l'enseignement, il est également utile d'examiner la théorie politique des systèmes d'organisation. Dans cette optique, les participants de l'organisation doivent être considérés comme des acteurs politiques ayant des besoins qui leur sont propres, des objectifs et des stratégies destinées à les atteindre. Leur réussite dépend de rapports de pouvoir qui peuvent être complexes et protéiformes, mettant en jeu à la fois l'autorité et l'influence. Par exemple, l'autorité peut être transférée jusqu'à un certain point en ayant recours au schéma de la prise de décision, et l'influence peut s'exercer par la concertation. L'influence peut aussi s'exercer au moyen des manipulations qui se passent « dans les coulisses » et déterminent les questions qui peuvent être traitées publiquement et celles qui sont occultées.

Les schémas de prise de décision et la concertation sont souvent utilisés aux États-Unis où le système éducatif est fortement influencé par une solide tradition de contrôle local (c'est-à-dire municipal) des écoles, auquel s'ajoutent la responsabilité juridique de l'enseignement qui se situe généralement au niveau de l'État, et un vif intérêt porté aux questions d'enseignement par l'administration fédérale. Qu'il s'agisse de l'État ou de l'administration fédérale, les schémas de décision leur donnent la possibilité d'exercer leur autorité sur des questions d'ordre général, tandis que les détails sont réglés à un niveau de décision inférieur. Par ailleurs, la concertation est, elle aussi, souvent utilisée, notamment entre l'école et le niveau local.

Il ressort d'une synthèse des théories que ce sont les capacités différentes à façonner, interpréter et manipuler l'environnement institutionnel qui déterminent en grande partie les modes d'organisation qui sont, ou ne sont pas, adoptés. Autrement dit, les individus et les coalitions favorisent les éléments qui sont conformes à leurs intérêts et tentent de supprimer ceux qui ne le sont pas. En outre, au fil du temps, certains changements d'organisation peuvent amener l'environnement institutionnel à évoluer et à changer.

Considérons, à titre d'exemple, le train récent de réformes visant à la décentralisation. Aux États-Unis, les mesures en faveur de la décentralisation s'inscrivent à la rubrique de la gestion de l'école par elle-même, et supposent le plus souvent que les établissements disposent d'une plus grande autorité en matière de budget, de recrutement ou de programme d'études. Cependant, la répartition de l'autorité au niveau de l'école (entre les chefs d'établissement, les enseignants ou les parents d'élèves) dépend de la capacité des individus ou des coalitions à mettre à profit certains aspects de l'environnement institutionnel, et notamment les valeurs telles que l'efficacité, le professionnalisme et la démocratisation. La gestion de l'enseignement par l'école fait partie des changements liés au mouvement de restructuration qui a commencé aux États-Unis au milieu des années 80. Bien que la restructuration ait suscité une grande activité de la part des décideurs, les contrôles institutionnels se sont montrés résistants et limitent la portée de la prise de décision par l'État, l'administration fédérale et l'école. Les modifications spectaculaires de l'autorité décisionnaire générale des écoles constituent l'exception plutôt que la règle.

Ces constatations permettent d'analyser en connaissance de cause les résultats d'enquêtes actuellement disponibles et de proposer le développement futur du question-

naire. Par exemple, les optiques théoriques exposées ci-dessus montrent qu'il importe d'examiner le mode aussi bien que le niveau de décision dans toute analyse de la décentralisation. Pour tous les pays confondus, seule la moitié environ des décisions examinées au titre de l'enquête de l'OCDE sont prises de façon autonome. Aux États-Unis, moins d'un cinquième de toutes les décisions sont autonomes. Qui plus est, on peut déduire de la théorie institutionnelle que les décisions relatives à l'organisation de l'enseignement sont généralement prises au niveau de l'école. Les résultats de l'enquête confirment généralement ce postulat.

Les propositions en vue d'un élargissement de l'enquête comprennent l'inclusion de certains modes informels d'influence; l'action des acteurs extra-scolaires tels que le public, les parents d'élèves et les syndicats; l'ampleur des schémas de décision; la mesure dans laquelle les conseils donnés au titre de la concertation sont suivis; les processus de décision qui agissent à l'intérieur de chaque niveau, ainsi que d'un niveau à l'autre, et le degré d'application des décisions. Il est aussi proposé d'ajouter à l'enquête les décisions concernant l'évaluation des enseignants, de l'école et des élèves.

*

* *

1. Introduction

In 1990, Network C of the OECD INES project identified school-level decision-making as a topic of interest in Member countries. Noting recent increases in decision-making authority allocated to schools, Network members set out to describe school decision-making in the context of the overall decision-making structure of their education systems. An instrument to collect information from experts – the Survey on Locus and Modes of Decision-Making – was piloted and subsequently revised, and resulting indicators focusing on school-level decision-making were included in the first and second editions of *Education at a Glance* (OECD, 1992 and 1993). Meuret (in OECD, 1995) presents an extended analysis of these data.

Since the survey was initially conceived, interest in decision-making in education systems has broadened, and discussions of the survey results have deepened network members' appreciation of the complexity of decision-making in educational organisations. For example, although a particular level may hold formal authority, decision-making often represents the combined activity of many levels – both formally and informally. In fact, while it may be the exception in some countries, decision-making shared between levels is very common in others. Also, although increases in school-level decision-making are commonly characterised as representing increased decentralisation, the distinction between decentralisation and centralisation must be applied with great care. Because decentralisation in one arena may be countered by centralisation in another, it must be recognised that there is no simple dichotomy between the two. Lastly, there is now increased interest in decision-making at levels other than the school – the country

and intermediate levels – and in the role of extra-organisational actors, such as the public and unions.

The purpose of this chapter is to introduce ideas that can serve as a beginning for building on the accomplishments of the initial survey. The chapter begins by discussing two areas of organisational theory, selected to provide a conceptual framework for schools as organisations and for decision-making itself. Although the framework spans a broader range of topics than the Network C survey, its concepts are useful for analysing the data and for considering revisions to the data collection. Institutional theory, presented first, notes that to understand how educational organisations operate, it is critical to recognise that decision-making is affected not only by explicit structures but also by established norms of behaviour and values, *i.e.* institutions. Conformity with these institutions is critical to maintaining public confidence and organisational legitimacy. Political systems theory, discussed next, focuses on the structure and process of decision-making itself. The term “political”, as employed here, encompasses all aspects of power relationships in governance. Thus, the term includes, but is not restricted to, the relationships of specific parties and governments. Through the ideas that both authority and influence shape decision-making, that the process operates in both a formal and informal arena, and that all this takes place within a broader social context, this perspective facilitates the development of descriptions that capture the complexity of decision-making. A brief discussion of decentralisation reforms illustrates how a synthesis of these approaches can contribute to the understanding of the decision-making process.

The second section of the chapter examines the Network C locus of decision-making survey. After presenting ideas for analysis of the existing data (describing decision-making in 1991), we offer initial suggestions for directions in which the survey could be expanded.

In both sections, the ideas presented are illustrated with examples primarily from the United States. The United States was chosen because it is the system about which we have the most knowledge, and because it is our starting point for exploring the usefulness of different theoretical perspectives and for generating ideas for analysis and expansion of the survey.

2. Theoretical Perspectives

Institutional Theory

Traditionally, formal organisations – such as schools and education systems – are understood to develop rational bureaucratic structures as a means of co-ordinating and controlling the relationships needed to execute work activities that utilise complex technologies (Scott, 1975). This concern for the impact of technology on organisational structure has its roots in classical organisational theory, which holds that the success or failure of organisations operating in markets is determined by their ability to produce outputs efficiently (Weber, 1946). To maximise efficiency in the production process, administrators are thought to focus much of their attention on problems arising in their organisation’s “core technology” (Thompson, 1967). As markets expand and as technol-

ogies become increasingly complex, organisations grow in administrative intensity, develop rationalised formal structures (e.g. bureaucratic offices and rules of operation), and increase the level of control over their technical activities.

While this conceptualisation of organisational structure and change may be generally applicable to organisations operating in markets where survival depends on efficiency – e.g. manufacturing organisations – it is less appropriate when applied to educational organisations. In educational organisations, instructional technologies are uncertain at best and, in the judgement of some observers, lacking entirely (Corwin, 1974; Jackson, 1968). Moreover, in most countries, public education systems are relatively sheltered from the economic forces of the marketplace.

Not surprisingly, researchers in the United States have found little evidence to suggest that the increasing complexity of instructional technologies has played a significant role in the expanding bureaucratisation of schools. For example, the number of administrators per US school district grew from 0.23 in 1932 (the earliest date for which such records are available) to 6.8 in 1970 (US Bureau of the Census, 1975), while teaching practices remained virtually unchanged (Dreeben, 1973). More complex technologies were introduced into the classroom beginning in the early 1970s, but with no observable impact on administrative intensity, formalisation, or frequency of evaluation (Cohen *et al.*, 1979; Davis *et al.*, 1976). Time allocation studies cast further doubt on the model that the increased complexity of technology has been the driving force behind the administrative growth of education systems. Both observational (Goodman, 1976; Sproull, 1976) and survey studies (Meyer and Rowan, 1978) have demonstrated that in the United States, only a small percentage of school administrators' time is devoted to "instructional leadership" tasks such as programme planning and development, staff evaluation, and monitoring.

In the light of these observations, theorists have begun to argue that "the history of schools has been misinterpreted as the emergence of organisations that co-ordinate the technical work of education" and have proposed that the structure of an education system arose, not so much to maximise efficiency, but "to bring the process of education under a socially standardised set of institutional categories" (Meyer *et al.*, 1981). According to this perspective, educational organisations are viewed as "institutional" organisations whose structural arrangements are shaped more by institutional rules than by the demands of production (Meyer and Rowan, 1977).

Drawing upon the work of Bidwell (1965) and Weick (1976), Meyer and Rowan (1978) have observed that education systems tend to shelter their organisational structures from the technical work activities, thus allowing instructional activities to be carried out with relatively few direct controls in the relative isolation of the classroom. This is made possible to a large extent by the establishment and regulation of categories – such as grade levels, subjects, courses, teachers, and students – that are regarded as legitimate by all constituencies involved. Thus, in educational organisations, "structure is disconnected from technical [work] activity and activity is disconnected from its effects" (Meyer and Rowan, 1978). This process serves two key functions. By avoiding close supervision of instructional activity and outputs, educational organisations 1) mask the inconsistencies and inefficiencies that occur in their core technologies, and 2) make local variation in the

content and effectiveness of instructional practices invisible. In the absence of technical controls, educators adopt a "logic of confidence", in which higher levels of the system organise on the assumption that lower levels are operating effectively as long as they conform to institutional rules. Thus, adoption of particular organisational structures that are viewed as rational and legitimate benefits school systems not because it increases efficiency, but rather because it leads to high levels of support and legitimacy.

The relatively low degree of implementation of policies concerning curricular matters provides convincing evidence of such "loose coupling" (Weick, 1976). Whether curricular policies are determined at the local level, as in the United States, or at the national level, as in Finland, France and Sweden, the official curriculum may be only loosely connected to what teachers actually teach in the classroom (Haenisch, 1993a; 1993b; Cohen *et al.*, 1990).

Further evidence of loose coupling in educational organisations comes in the form of their characteristically weak controls over instructional methods. In the United States, Rowan (1990) found that more than three-quarters of a nationally representative sample of high-school teachers reported that they were observed only once during the school year. In many European countries, the existence of "inspectors" would seem to ensure tighter controls over teaching practice. However, a recent study by Hopes (1991), summarising the results of a two-year project reviewing the work of inspectors in each of the twelve European Community school systems, casts doubt on this assumption. Inspections typically take place only once every few years and, in some countries (e.g. Denmark, Greece, Italy and Portugal), "inspectors" are limited to an information-gathering role that does not include the actual evaluation of teachers' performance. In countries where evaluations of teachers are conducted, they are often not systematically carried out and, according to Hopes, have become little more than "a sporadic ritual". Although they found an association between state (country-level) control over curriculum and the extent to which the same mathematics curriculum is taught by teachers, Stevenson and Baker (1991) note that regardless of the degree of centralisation, classroom instruction is influenced by a multitude of factors, many of which operate at the school level.

Finally, the loose coupling of educational organisations is demonstrated by their lack of control of work through regular inspection of outputs. Although many education systems routinely make use of tests to assess and determine the life chances of students, such measures are rarely used to evaluate the performance of teachers or schools. Cross-sectional testing programmes in the United States and national examinations of certification employed in many European countries impose some measure of "accountability", but they are limited because of their emphasis on absolute outputs rather than on "value added". One exception to this rule is the French system of student evaluation, which employs both absolute and "value added" measures, based on representative samples of students that are monitored throughout their school career (OECD, 1991). Generally, however, in European countries, because schools "are assumed to be operating optimally and satisfactorily through compliance with predetermined standards" (Hopes, 1991), inspectorates rarely evaluate their performance.

Such analyses illustrate the contribution of institutional theory to a broad understanding of how education systems work: recognition of the importance of established

norms and practices in both the day-to-day operation and in the decision-making process in education systems. However, the theory focuses on general principles; it is relatively unconcerned with actually describing how particular educational organisations operate or how they change, specifically the structure and process of decision-making. For answers to these questions, we turn to a discussion of educational organisations as political systems.

Educational Organisations as Political Systems

To understand the development and functioning of decision-making structures, it is useful to look at political systems theory of organisations. According to this perspective:

- Organisations are best conceived of as political systems, both internally and in their external relationships;
- Organisational participants should be viewed as political actors with their own needs, objectives, and strategies to achieve those objectives;
- Coalitions of actors emerge in organisations, identify collective objectives, and devise strategies to achieve their objectives; and
- Both individual and collective actions are constrained by organisational structures, technologies, and ideologies (Bacharach and Lawler, 1980). (Elements of a political theory of organisations can also be found in the works of Gouldner [1954] and Cyert and March [1963].)

Applying this framework to education systems, the best approach for understanding the decision-making process is to examine the actions of individuals and groups – both within and outside the system – as they attempt to establish structural arrangements for decision-making, and as they seek to achieve their objectives within those structures.

The identification of power relationships and the associated ability to influence decisions are essential to understanding schools as political systems. While power relationships in organisational systems are multi-faceted, they can be broadly subsumed under the concepts of authority and influence (Bacharach, 1978; Bacharach and Aiken, 1976). Authority refers to the legitimate decision-making power that resides in the various positions in organisational systems, whereas influence refers to the informal exercise of power to guide decision-makers (Simon, 1950; Bierstadt, 1950). Although authority relationships are limited and coterminous with organisational boundaries, influence networks are unlimited and may extend beyond the organisation into the environment.

A fundamental and recurring dilemma in the analysis of organisational decision-making is the failure to explicate the distinction between authority and influence processes. In fact, the absence of authority in no way means that power is also absent. Thus, Bacharach (1981) observes that "much of the debate over whether organisations are centralised or decentralised has been theoretically circumscribed because it is never clear whether organisational theorists are directing their attention to authority or influence".

When power takes the form of authority, its use is confounded by two potentially conflicting imperatives. On the one hand, decision-makers with authority may realise that

good decision-making is dependent on high-quality information which they may not have. On the other, these decision-makers are reluctant to relinquish their authority. This dilemma is commonly addressed in two ways: 1) through the establishment of decision-making frameworks that set parameters within which others can make decisions and 2) through a process of consultation. The extent to which authority relationships are modified through the use of these mechanisms in the decision-making process depends upon how they are used. Where decision-making frameworks are relatively broad or where guidance obtained through consultation is almost invariably followed, formal authority is all but delegated to subordinate levels. But, where decision-making frameworks are narrow or where recommendations have only a marginal effect on policy decisions, these mechanisms may provide little more than the illusion of input to the process.

Both mechanisms are used quite commonly in the United States, where the education system is greatly influenced by a strong tradition of local (*i.e.* municipality-level) control of schools, coupled with statutory responsibility for education generally residing at the state level. Because local or school decision-making is almost always viewed as preferable, decision-making at the state level – especially in controversial matters – is often seen as intrusive. Also, the Federal government does not have direct authority over education; its involvement stems from its role in ensuring equal opportunity, civil rights, and general national well-being. For both the state and Federal levels, decision-making frameworks offer an opportunity to exercise authority over general matters while allowing details to be specified at a lower decision-making level.

Consultation is also frequently used in decision-making – particularly between the school and district (local) levels. Consultations may be initiated either by the district or the individual schools themselves. Decisions made by the district in consultation with schools are often guided by a task force or committee comprising administrators, teachers and other potential local players such as parents and community members. The task force investigates the issue and makes recommendations to the district which are generally not binding. School principals also provide input on decisions through meetings called by district-level staff. Also, it is a common practice for individual schools to raise questions of concern informally for the school district's response. For example, issues may be brought to the attention of school administrators at school staff meetings and then passed on to the district by the school principal, either verbally or in writing.

Influence relationships may occur both within organisations and between organisations and their environments. Perhaps the most potent use of influence occurs in "behind the scenes" manipulations that determine which issues are addressed publicly and which issues are suppressed. Bachrach and Baratz (1962) argue that there are, in fact, two "faces of power" – one manifest in political conflict and its resolution, the other expressed covertly through the ability of powerful interests to limit the agenda of decision-making to relatively safe issues. Thus, issues can be suppressed through non-public deliberations or, even more effectively, through the creation by powerful interests of a "mobilisation of bias" (*i.e.* widespread and pervasive values and beliefs) which delimits what it is safe to do and what it is not.

Within the boundaries set by the mobilisation of bias, there exists a "zone of tolerance" within which authority may be exercised in a routine manner (Boyd, 1979). However, when those in authority exceed these invisible boundaries – which may be broad or narrow – they run the risk of coming into conflict with community or broader social values and encountering widespread controversy and opposition. In the case of the United States, the relative vulnerability of public school administrators makes them particularly cautious about testing the limits of their authority. Unbridled conflict often results in the dismissal of the district superintendent, and, if controversy persists unchecked, it can challenge the very legitimacy of the school system (Bacharach, 1981). Conflict avoidance thus becomes a salient orientation of school administrators and, as such, places real constraints on decision-making.

In cases when conflictual issues do enter into the public arena, influence is exercised through the process of negotiation. Parental groups may negotiate for greater representation in decision-making, teachers' unions may press for better working conditions or smaller classes, and business interests may seek to influence curricular decisions. Though decision-making by negotiation is often characterised as "consensus building", the threat of conflict always lurks just below the surface of public debate, and public manifestations of controversy – demonstrations by parents, teachers' strikes, etc. – often signal the breakdown of negotiations.

Synthesis of the Perspectives

The political perspective outlined above offers significant insights into the particular characteristics of different educational organisations and the processes by which they change. Through the exercise of authority and influence, actors within and beyond organisational boundaries exert substantial control over decision-making processes and, hence, determine the direction of organisational change. Such control is exercised both directly (by invoking authority of office) and indirectly (by influencing those in authority or through the "mobilisation of bias" in favour of specific organisational forms).

In its present form, institutional theory offers its own insights into the importance of successfully institutionalised organisational forms and practices, yet it tells us very little about "institutionalisation" as a process – *i.e.* where institutions come from, why some structured innovations spread while others do not, and how decision-making processes influence change (DiMaggio, 1988). In large part, institutional theorists have succeeded in excluding agency and interest from its formulations by focusing on institutionalisation as an "achieved state". When viewed as a *process*, however, institutionalisation is profoundly political, reflecting the relative power of organised interests as well as the actors who mobilise around them (DiMaggio, 1988). Recalling Gouldner's (1954) words, institutions have never "developed and operated without the intervention of interested groups" and an institution's persistence is often the "outcome of a contest between those who want it and those who do not".

Although a small group of organisational theorists have recently begun to acknowledge the role of agency and interest in explaining institutional processes, the precise manner in which individuals and coalitions influence institutional forms has not been clearly articulated. Political theories of organisational change, with their emphasis on the

"mobilisation of bias", provide critical clues to the process of institutional metamorphosis, but these conceptual tools have not been successfully incorporated into institutional theory. Hence, we are reminded that interests are institutionally defined and shaped (Friedland and Alford, 1987), while less attention is given to the fact that those who can shape the institutional environment possess an important form of power.

DiMaggio (1988) has observed that institutional forms do not "arise like mist from the environment", but rather, represent the compelling interests of various actors. Extending this argument, we suggest that it is the differential ability of divergent interests to shape, interpret, and manipulate the institutional environment that, in large part, determines which organisational forms are adopted and which are not. That is, individuals and coalitions – internal and external to the organisation – will promote elements which are consistent with their interests and attempt to suppress those which are not.

By way of illustration, consider broadly the recent wave of reforms aimed at decentralisation. Meuret (OECD, 1995) has suggested that, while in some cases, such reforms are motivated by political goals of democratisation, in other instances, they are promoted by those seeking to increase the efficiency of the system. The ideas outlined above offer insight into the sources of such differentiation.

In the United States, decentralisation reforms have been introduced under the rubric of school-based management, in which increased decision-making authority is granted to individual schools, typically in the areas of budgeting, staffing, or curriculum. While these plans share certain common characteristics, they also differ greatly in the manner in which authority is distributed within the school level. Researchers have taken note of the fact that schools operating under school-based management vary substantially in terms of the relative influence that local constituencies bring to bear on the decision-making process (Wohlstetter and Buffett, 1991; Wohlstetter and McCurdy, 1991). As a means of characterising this relative "balance of power" at the school level, three models of school-based management have been proposed: the principal-control model, the teacher-control model, and the community-control model (Wohlstetter and Odden, 1992; Drury and Levin, 1994).

Which of these three models arises can perhaps be best explained in terms of the ability of individuals and coalitions to define, shape, and manipulate the institutional environment to further their own interests. In districts where central office administrators have provided the impetus for reform, school principals are often given the bulk of decision-making authority at the school level and institutional values such as efficiency and educational effectiveness typically prevail. In districts where teachers have organised themselves into strong collective bargaining units, reform more often emphasises teacher empowerment in support of professionalisation goals. Finally, in districts where community interests are actively engaged in school reform efforts, school-based management is likely to result in the empowerment of parents, served up under the mantle of democratisation.

The current debate over ability grouping in the United States provides another illustration of the manipulation of the institutional environment to support particular interests. According to Slavin (1990), arguments in favour of ability grouping tend to focus on *effectiveness*, while arguments opposed to grouping typically focus at least as

much on *equity* as on effectiveness and on *democratic values* as much as on outcomes. Which of these ideological elements achieves greater salience in the institutional environment will largely determine which structural form is adopted. In turn, the relative primacy of various environmental elements (*i.e.* ideologies, myths, and taken-for-granted understandings) will be determined by the differential influence of competing interests.

School-based management is one of a number of changes in decision-making associated with the restructuring movement in the United States, which began during the mid-1980s. In addition to school-based management, restructuring has generated a sizeable level of policy activity at the state level, such as increases in required achievement testing programmes. Further, the popularity of restructuring as a school-improvement strategy has contributed to acceptance of Federal initiatives in the area of achievement testing, standards, and national goals. In contrast to school-based management, the state and Federal level reforms are seen as centralising forces. However, institutional controls limiting the scope of state and Federal decision-making are resilient, and the bulk of the energy devoted to the restructuring has gone into fine-tuning the decision-making relationships between schools and districts (such as giving schools increased latitude in decision-making frameworks), and introducing changes within the school level. Dramatic changes in the overall decision-making system have been the exception rather than the rule.

Similar explanations have been proposed for the structure of recent reforms in other OECD countries. Depending upon the relative influence exercised by business, professional, and community interests, decentralisation reforms are likely to emphasise different aspects of the institutional environment and manifest themselves quite differently. Thus, in countries suffering recent economic setbacks, such as Australia, business interests promote the goal of efficiency, and restructuring takes the form of "administrative decentralisation", in which authority is devolved from superior to subordinate organisational units (Chapman *et al.*, 1993). In other countries, such as Spain, local community interests promote the goal of democratisation, and restructuring is more often in the form of "political decentralisation", where authority is transferred from central to regional authorities or, at the school level, to parents (Muñoz-Repiso *et al.*, 1993).

3. The Network C Locus of Decision-Making Survey

The institutional and political approaches presented above describe educational decision-making in the broadest terms, extending well beyond the practical concerns of the present Network C survey. In fact, some aspects of the theories will probably never be appropriate for a survey of this type. Nonetheless, these theoretical perspectives provide a useful basis for analysing current survey data and for proposing future expansions of the survey instrument. Eventually, they could also serve as a starting point for specifying a more detailed conceptual framework to guide the development of the Network C survey and to place it in the context of a broader theory of decision-making.

The Current Survey

The current OECD survey consists of 35 questions concerning the locus of decision-making in four broad "fields":

- organisation of instruction;
- planning and structures;
- personnel management; and
- resources.

In addition, four "levels" of decision-making are identified. (At each level, only those participants with formal authority are considered; the influence of parents' associations, teachers' unions and other interest groups is beyond the purview of the current survey):

- the school;
- the inferior intermediate level (Intermediate 1);
- the superior intermediate level (Intermediate 2); and
- the country.

Finally, three "modes" of decision-making are distinguished:

- fully autonomous;
- in consultation with bodies located at another level; or
- within a framework set by a higher authority.

It is important to note that, in terms of this survey, decisions made in "consultation" do not include those made in conjunction with organisational participants at the same level, nor do they include decisions reached in conjunction with actors outside of the educational bureaucracy.

If this framework is applied, the decision-making process in each of fourteen participating countries may be characterised separately for ISCED levels 1, 2, and 3 and for public and private education. In the present context, we shall focus exclusively on lower secondary public education (ISCED 2) in the fourteen participating countries. The data presented here are drawn from Meuret's summary of 1991 OECD survey data and from *Education at a Glance* (OECD, 1993).

Level of Decision-Making

For all fourteen countries combined, decisions at the lower secondary level tend to be relatively decentralised: 38 per cent at the school level; 31 per cent at the inferior intermediate level; 12 per cent at the superior intermediate level; and 19 per cent at the country level. Disaggregated by country, less than 10 per cent of the 35 decisions considered in the 1991 survey are made at the national level in Belgium, Germany, Sweden, Switzerland and the United States, and less than half are made at the national level in Austria, Denmark, Finland, France, Ireland, New Zealand, Norway, and Spain. Portugal is the only country where a majority of decisions are made at the national level.

Mode of Decision-Making

While data on the level of decision-making provide an overall picture of the relative centralisation of educational organisations in the 14 participating countries, the theoretical framework developed earlier in this chapter underlines the importance of examining the *mode*, as well as the locus of decision-making. To assess accurately the degree of centralisation in education systems, it is as important to know the extent to which authority is not autonomous and which levels influence decisions, as it is to identify the levels where decisions are ultimately made. For all countries combined, only about one-half of the decisions examined in the OECD survey are made autonomously, about 15 per cent are made in consultation with another level, and approximately one-third are made within a decision-making framework established at a higher level.

In this regard, the decision-making process in the United States is clearly distinguished from that of other countries. In nine countries, the majority of decisions are made autonomously and, in four countries, more than a third of decisions are reached in this manner, but in the United States, less than one-fifth of all decisions are made autonomously. The remainder are about evenly split between decisions made in consultation with another level and those made within a set framework. Because these modes are used so frequently, it is important to incorporate this dimension into analyses of decision-making in the United States.

When decisions made at the school level are under examination, this distinguishing characteristic of American education is especially critical. In terms of autonomous decisions, US schools rank close to the bottom among the fourteen participating countries. However, when one includes decisions made at the school level in consultation with another level, decisions made within a decision-making framework established by a higher level, and decisions reached in consultation with schools, US schools rank fourth (tied with Denmark and Finland) in terms of overall participation in decision-making. These results are consistent with the discussion above about the use of decision-making frameworks and consultation in US schools.

Other countries also rank quite differently depending upon which criterion is employed. For example, Belgian schools rank second in terms of autonomous decision-making, but next to last in terms of overall participation in decision-making.

Field of Decision-Making

In addition to the locus and mode of decision-making, it is important to consider the specific "field" in which decisions are made. Institutional theory suggests that the core technology of educational organisations tends to be only loosely coupled with other organisational elements. If this is the case, one would expect decisions regarding the organisation of instruction to be made largely at the school level, often in the confines of individual classrooms. The survey data generally support this view. In thirteen of the fourteen countries considered, the majority of decisions concerning the organisation of instruction are made at the school level (Switzerland being the sole exception). However, school-level decisions are often made within a framework established at a higher level. To the extent that decision-making frameworks are narrow – restricting the range of alternatives at the school level – a degree of authority remains at the higher level.

Unfortunately, the 1991 survey data do not permit an analysis of the latitude of decision-making frameworks. Nor do these data indicate where, within schools, such decisions are made. Separate survey data show that more than 80 per cent of teachers believe they have considerable influence over selecting teaching techniques, evaluating and grading students, and determining the amount of homework to assign (US Department of Education, 1994).

Extending the Network C Survey

The foregoing discussion of the Network C survey data is wholly consistent with the theoretical perspectives described in this chapter. However, in its present form, the survey does not address many other aspects of decision-making in education that are suggested by the theories. Hence, we conclude this section by focusing on those areas in which the survey could be extended and refined.

Level of Decision-Making

Currently, the survey of educational decision-making focuses exclusively on the locus of *formal decision-making arrangements* in education systems. Earlier, however, we stressed the importance of examining the complexities of power relationships. In this regard, Bacharach (1981) observed that much of the debate over the relative centralisation or decentralisation of organisations is due to a lack of clarity concerning the type of decision-making power under consideration. Meuret (in OECD, 1995) clearly recognises this as an area in which the present survey instrument could be expanded when he observes that the influence and power of a level within the education system extends beyond the number of decisions it makes. As an illustration, he states that because of activities such as publicising test results and specifying broad goals, the US Department of Education has an influence on the American system which goes beyond that which is measured in the survey (OECD, 1995). This concern could be addressed most directly by extending the current survey to include those decisions involving the exercise of informal influence as well as those which rely upon the exercise of formal authority.

While the current survey identifies the level at which a decision is made, it does not single out the particular actor or coalition involved in decision-making within that level. For example, it does not distinguish between elected officials and civil servants, or between administrators and teachers. Related to this is the survey's exclusive focus on intra-organisational actors. Parents' groups, teachers' unions, and other interest groups are explicitly excluded from consideration, though their influence may be substantial. These issues become especially important in determining the extent to which the objectives of decentralisation projects are achieved. For example, where teacher professionalisation is a central goal, it would be useful to know whether or not teachers actually participate in the decision-making process. Similarly, where democratisation is emphasised, it is important to know the degree to which parental and community interests are involved in decision-making. Adding a lateral dimension, providing information on the relative authority or influence of administrators, teachers, parents and other stakeholders, would enhance the survey in this respect.

Mode of Decision-Making

In its present form, the survey distinguishes between three "modes" of formal decision-making: autonomous decisions; decisions made in consultation with another level; and decisions made within a framework established at a higher level.

Yet decisions reached through consultation may be nearly autonomous, while others may involve a real sharing of authority. Similarly, some decision-making frameworks may be quite narrow, while others may be relatively broad and unrestrictive. Inter-coder reliability and the validity of the resulting data would be enhanced considerably if such distinctions were made explicit. For example, decision-making frameworks might be characterised as highly restrictive, moderately restrictive, and unrestrictive, and consultation might be described in terms of the relative influence of consulting and consulted levels. Meuret (in OECD, 1995) notes, that, in completing the 1991 survey, respondents sometimes exhibited confusion and uncertainty in assessing the appropriate mode of decision-making. Where a decision was reached after *within*-level consultation, some countries characterised the decision as autonomous (a correct response in terms of the current survey), while others did not reply at all. Others displayed confusion in cases where decisions reached through consultation seemed to be influenced less by the level *seeking* consultation than by the level *consulted*. Finally, countries sometimes displayed uncertainty in distinguishing between decisions made autonomously and those made within a framework established at a higher level. Where frameworks were relatively unrestrictive, for example, decisions were occasionally characterised as autonomous.

The current survey might also be expanded to include questions concerning the *implementation* of decisions. At whatever level decisions are made, if they are not implemented, they cannot influence policy. In the light of our emphasis on the institutional nature of educational organisations and the attendant "loose coupling" of policy and technical work activities, it would seem critical to evaluate the degree to which decisions are actually carried out.

Finally, because the present survey explicitly excludes informal influence relationships, the modes of decision-making considered are limited to those which characterise formally recognised relationships. However, informal modes of decision-making are often equally important, both within educational organisations and between educational organisations and their environments. In the United States, for example, superintendents sometimes exercise informal influence in excess of their formal authority by controlling the flow of information and setting policy agendas. Also, external actors – such as business coalitions and teachers' unions – often exercise substantial influence over decision-making in educational organisations, even though they have no formal authority at all. The present survey could thus benefit by extending the present modes of decision-making to include such informal mechanisms as negotiation and the establishment of zones of tolerance and by including external, as well as internal, actors.

Field of Decision-Making

Presently, the survey addresses four fields of decision-making: organisation of instruction; planning and structures; personnel management; and resources. Although the fact that the majority of decisions concerning the organisation of instruction are made at

the school level is consistent with our interpretation of educational organisations as institutional organisations, other fields relevant to our conceptual framework are not addressed. It is especially important to know where decisions concerning the evaluation of teachers, schools, and student outputs are made and how faithfully they are implemented.

4. Conclusion

Since 1990, when Network C of the OECD INES project identified school-level decision-making as the topic for a survey, interest in decision-making structures has broadened, and discussions of the results have deepened members' appreciation of the complexity of decision-making in education systems. The purpose of this chapter is to introduce ideas that can serve as a beginning for building on the accomplishments of the original survey.

First, to provide a conceptual framework for decision-making and for schools as organisations, two theoretical perspectives have been presented. Institutional theory notes that the regulation of categories (such as grade levels, subjects, courses, teachers, and students) which are highly associated with institutionalised norms of behaviour and values allows education systems to shelter their technical work activity (instruction of students) from organisational controls. It is assumed that the system is operating effectively as long as its activities conform to these institutional norms and categories. In addition, the behaviour of organisational actors is affected as much by institutional expectations as by policies developed through the decision-making process. Political systems theory addresses the issue of the decision-making process, in which power relationships are multi-faceted and complex. Some authority may be transferred through the use of decision-making frameworks, and influence may be exercised through the consultative process. A synthesis of the theories suggests that the differential ability of divergent interests to manipulate the institutional environment determines which organisational forms are adopted.

The concepts from these theories are useful for analysing decision-making in education systems. To use the United States as an example, a strong tradition of local control, coupled with statutory responsibility for education residing at the state level and a strong Federal interest in educational matters, has resulted in widespread use of decision-making frameworks and consultation. In recent school-based management reforms, allocation of authority within the school level (to principals, teachers, or parents) has been affected by the ability of individuals and coalitions to draw upon aspects of the institutional environment, including the values of efficiency, professionalisation, and democratisation. Although restructuring has generated a great deal of policy activity in the United States, institutional controls have been resilient, limiting the scope of state, Federal, and school decision-making.

These theoretical approaches also provide a useful basis for analysing current survey data and for proposing future expansions of the survey instrument. On the average, only about half the decisions are made autonomously (in the United States, the figure is less than one-fifth), suggesting that results regarding the mode of decision-making should be

analysed carefully. Suggestions for expanding the survey include incorporating informal influence processes; extra-organisational actors such as the public, parents, and unions; the breadth of decision-making frameworks; the extent to which advice given in consultation is followed; decision-making processes within as well as across levels; and the degree of implementation of decisions. It is also proposed that decisions regarding evaluation of teachers, schools, and students be added to the survey.

Acknowledgement

The authors would like to thank the National Center for Education Statistics in the United States Department of Education for supporting the preparation of this chapter, and David Baker, Dawn Nelson, and Renée Sherman for their helpful comments on earlier drafts.

References

- BACHARACH, S. (1978), "Morphologie et processus: une critique de la recherche intra-organisationnelle contemporaine", *Revue de Sociologie du Travail*, Vol. 20, pp. 153-173.
- BACHARACH, S. (ed.), (1981), *Organizational Behavior in Schools and School Districts*, Praeger Publishers, New York.
- BACHARACH, S., and AIKEN, M. (1976), "Structural and process constraints of influence in organizations: a level-specific analysis", *Administrative Science Quarterly*, Vol. 21, pp. 623-642.
- BACHARACH, S.B., and LAWLER, E.J. (1980), *Power, Coalitions, and Bargaining: The Social Psychology of Organizational Politics*, Jossey-Bass, San Francisco, CA.
- BACHARACH, P., and BARATZ, M.S. (1962), "Two faces of power", *American Political Science Review*, No. 56, pp. 947-952.
- BIDWELL, C.E. (1965), "The school as a formal organization", in J. G. March (ed.), *Handbook of Organizations*, Rand McNally, Chicago.
- BIERSTADT, R. (1950), "An analysis of social power", *American Sociological Review*, Vol. 15, pp. 730-738.
- BOYD, W.L. (1979), "The politics of curriculum change and stability", in J. Baldrige and T. Deal (eds.), *The Dynamics of Organizational Change in Education*, McCutchan Publishing Corporation, Berkeley, CA.
- COHEN, E.G., DEAL, T.E., MEYER, J.W., and SCOTT, W.R. (1979), "Technology and teaming in the elementary school", *Sociology of Education*, Vol. 52, pp. 20-33.
- COHEN, D.K., PETERSON, P.L., WILSON, S., BALL, D., PUTNAM, R., PRAWAT, R., HEATON, R., REMILLARD, J., and WIEMERS, N. (1990), *Effect of State-Level Reform of Elementary School Mathematics Curriculum on Classroom Practice*, US Department of Education, Office of Educational Research and Improvement, Washington, DC.
- CORWIN, R.G. (1974), "Models of educational organizations", in F.N. Kerlinger and J.B. Carroll (eds.), *Review of Research on Education*, F.E. Peacock Publishers, Inc, Itasca, IL.
- CYERT, R.M., and MARCH, J.G. (1963), *A Behavioral Theory of the Firm*, Prentice Hall, Englewood Cliffs, NJ.
- DAVIS, M.E., ROWAN, B., and STACKHOUSE, E. (1976), "Loose and tight coupling in educational organizations", in M. Davis et al. (eds.), *The Structure of Educational Systems: Explorations on the Theory of Loosely Coupled Systems*, Stanford Center for Research and Development on Teaching, Stanford, CA.
- DIMAGGIO, P. (1988), "Interest and agency in institutional theory", in L. Zucker (ed.), *Patterns and Organizations*, Ballinger, Cambridge, MA.

- DREEBEN, R. (1973), "The school as a workplace", in R.M.W. Travers (ed.), *Second Handbook of Research on Teaching*, Rand McNally, Chicago.
- DRURY, D. and LEVIN, D. (1994), *School-Based Management: The Changing Locus of Control in American Public Education*, US Department of Education, Office of Research, Washington, DC.
- FRIEDLAND, R. and ALFORD, R. (1987), *Bringing Society Back in Symbols, Structures and Institutional Contradiction*, Paper presented at a Conference on Institutional Change, Center for Advanced Study in the Behavioral Sciences, Stanford, CA, May.
- GOODMAN, S.M. (1976), *The Principal's Leadership Role: Practice vs. Preference*, Montgomery County Schools, Rockville, MD.
- GOULDNER, A. (1954), *Patterns of Industrial Bureaucracy*, The Free Press, New York.
- HAENISCH, H. (1993a), "Change in the primary school", Part of OECD Report *The Effectiveness of Schooling and of Educational Resource Management*, Landesinstitut für Schule und Weiterbildung, Soest.
- HAENISCH, H. (1993b), "Translating guidelines and syllabi into action", Part of OECD Report *The Effectiveness of Schooling and of Educational Resource Management*, Landesinstitut für Schule und Weiterbildung, Soest.
- JACKSON, P.W. (1968), *Life in Classrooms*, Holt, Rinehart and Winston, New York.
- MEYER, J.W. and ROWAN, B. (1977), "Institutional organizations: formal structure as myth and ceremony", *American Journal of Sociology*, Vol. 83, pp. 340-363.
- MEYER, J.W. and ROWAN, B. (1978), "The structure of educational organizations", in M.W. Meyer et al. (eds.), *Environments and Organizations*, Jossey-Bass, San Francisco, CA.
- MEYER, J.W., SCOTT, W.R., and DEAL, T. (1981), "Institutional and technical sources of organizational structure: explaining the structure of educational organizations", in H. Stein (ed.), *Organization and the Human Services: Cross-Disciplinary Reflections*, Temple University Press, Philadelphia, pp. 151-179.
- MUÑOZ-REPISO, M. et al. (1993), *School Effectiveness and Educational Resource Management in Spain*, Centre of Research, Evaluation, and Documentation, Madrid.
- OECD (1989), *The Effectiveness of Schooling and of Educational Resource Management, Country Reports - France*, Paris.
- OECD (1991), "The role of the inspectorate and inspectors in the development and monitoring of school management and effectiveness", in *The Effectiveness of Schooling and of Educational Resource Management*, Paris.
- OECD (1992), *Education at a Glance: OECD Indicators* (bilingual), CERI, Paris.
- OECD (1993a), *Education at a Glance: OECD Indicators* (bilingual), CERI, Paris.
- OECD (1993b), *Aiming for Effectiveness in Australian Schooling*, Paris.
- OECD (1995), *Decision-Making Processes in the Education Systems of Fourteen OECD Countries*, Paris.
- ROWAN, B. (1990), "Commitment and control: alternative strategies for the organizational design of schools", in C. Cazden (ed.), *Review of Research in Education*, American Educational Research Association, Washington, DC, Vol. 16, pp. 353-389.
- SCOTT, W. R. (1975), "Organizational structure", *Annual Review of Sociology*, Vol. 1, pp. 1-20.
- SELZNICK, P. (1949), *TVA and the Grass Roots*, University of California Press, Berkeley, CA.

- SIMON, G. (1950), "Notes on the observation and measurement of political power", *Journal of Politics*, Vol. 15, pp. 500-516.
- SLAVIN, R. (1990), "Achievement of effects of ability grouping in secondary schools: a best-evidence synthesis", *Review of Educational Research*, Vol. 60 (3), pp. 471-499.
- SPROULL, L.S. (1976), *Managerial Attention in New Education Programs: Some Preliminary Findings*, Paper presented at the Conference on Educational Organizations as Loosely Coupled Systems, Palo Alto, CA.
- STEVENSON, D.L. and BAKER, D.P. (1991), "State control of the curriculum and classroom instruction", *Sociology of Education*, Vol. 64, pp. 1-10.
- THOMPSON, J.D. (1967), *Organizations in Action*, McGraw Hill, New York.
- US Bureau of the Census (1975), *Historical Statistics of the United States*, US Bureau of the Census, Washington, DC.
- US Department of Education (1994, August), *Who's in Charge? Teachers' Views on Control Over School Policy and Classroom Practices*, US Department of Education, Office of Educational Research and Improvement, Washington, DC.
- WEBER, M. (1946), *Essays in Sociology*, Oxford University Press, New York.
- WEICK, K.E. (1976), "Educational organizations as loosely coupled systems", *Administrative Science Quarterly*, Vol. 21, pp. 1-19.
- WOHLSTETTER, P. and BUFFETT, T. (1991), *School-Based Management in Big City Districts: Are Dollars Decentralized Too?*, University of Southern California, Center for Research in Education Finance, Los Angeles.
- WOHLSTETTER, P. and MCCURDY, K. (1991), "The link between school decentralization and school politics", *Urban Education*, Vol. 25, pp. 391-414.
- WOHLSTETTER, P. and ODDEN, T. (1992), "Rethinking school-based management policy and research", *Educational Administration Quarterly*, Vol. 28, pp. 529-549.

**La mise en place d'un dispositif d'indicateurs pour le pilotage
des établissements secondaires français**
*Implementation of a System of Indicators for the Steering
of French Secondary School*

par

Jean-Claude Emin

Département de l'évaluation des systèmes et des unités d'éducation
Direction de l'évaluation et de la prospective, Ministère de l'Éducation nationale
Paris, France

Le ministère français de l'Éducation nationale met en place un dispositif d'indicateurs destiné à tous les établissements scolaires du second degré, publics et privés sous contrat, soit quelque 12 000 établissements. Deux principes méthodologiques ont guidé sa conception. Ce dispositif permet à l'établissement, à la fois de se rendre compte de son fonctionnement et d'en rendre compte : c'est d'abord un outil destiné au chef d'établissement et à son équipe, un outil de diagnostic de la situation et du fonctionnement de l'établissement, d'évaluation de la politique qui y est conduite et de ses résultats, de régulation et de suivi de son projet ; mais c'est également un outil de compte rendu de son activité et de ses résultats.

Ce dispositif d'indicateurs permet une évaluation fondée, à la fois sur une batterie standard d'indicateurs relativement complète et commune à l'ensemble des établissements et sur des indicateurs spécifiques à chaque établissement et à son projet : la batterie d'indicateurs a été construite en faisant très largement appel à l'expérience et aux besoins des praticiens, tout en utilisant comme toile de fond les travaux de recherche sur l'efficacité de l'école. Dans la base standard, plusieurs indicateurs s'efforcent de dépasser l'appréciation des résultats bruts des établissements pour évaluer leur « valeur ajoutée » ; cette base standard comprend des indicateurs qui donnent une mesure de l'efficacité de l'établissement, c'est-à-dire de ce qu'il apporte à ses élèves (et de ce que ceux-ci tirent de cet apport) en éliminant, autant que possible, l'incidence de facteurs sur lesquels l'efficacité de l'établissement ne saurait jouer. L'ensemble de ce dispositif, expérimenté en fin

d'année scolaire 1993/94 dans près de 600 établissements, sera disponible dans tous les collèges et lycées fin 1995.

La mise en place de ce dispositif présente quelques caractéristiques importantes qui résultent de choix politiques et méthodologiques; on les exposera après avoir défini les principaux éléments du contexte dans lequel se situe l'opération et rappelé les conditions – réunies depuis peu – qui rendent possible le projet. Ce n'est qu'ensuite que les indicateurs dont tous les établissements de France devraient pouvoir disposer en 1995 seront présentés.

Summary

The French Ministry of Education is introducing a system of indicators for use by all secondary schools, both public sector schools and private ones under contract, some 12 000 institutions in all.

The system is meant to be of service and assistance to schools, which will be provided with a set of indicators constructed on the basis of data gathered by national information and management systems, together with a series of reference indicators calculated for the country as a whole and for the different academies so that schools can see how they compare with others; they will also be supplied with the means for calculating, tracking and keeping records of any specific indicators they themselves might wish to develop. The design of the system has been based on two methodological principles.

This system will enable the school to gauge and give an account of its own operation; it is primarily intended as an instrument for use by the head of the school and his or her team to diagnose the situation and operation of the school, to assess its policy and the results thereof, and to monitor and fine-tune its project. But it is also a means of reporting on the school's activity and performance. Colleges and lycées are public service agencies of the Ministry and as such are accountable not only of course to their supervisory authority, but also to the community at large.

In addition, the system will permit an evaluation based both on a relatively complete standard set of indicators, common to all institutions, and on indicators specific to each institution and its project. Practitioners' experience and needs have figured very largely in the choice of indicators. This was also based on research work on school effectiveness, which gives the system an essential degree of homogeneity. It is important, however, for each school to be able to go beyond this common core of indicators and describe and evaluate particular aspects of its operation or performance: in short, to create its own

specific indicators, providing it has or is able to gear itself up to collect the requisite information.

Several of the common core indicators are intended to evaluate not just schools' raw results but also the «value added» that they generate, thereby measuring the school's effectiveness in terms of what it actually imparts to its pupils (and what the pupils derive from this) by eliminating as far as possible the influence of factors that are extraneous to the effectiveness of the school, such as differences in the level of attainment of children entering the school. It is only on this basis that valid comparisons of schools' performance can be made.

The entire system, which was tested during the latter part of 1993/94 in nearly 600 schools, will be available to all colleges and lycées by late 1995. Having comprehensive data of this kind on the operation of all colleges and lycées should do much to improve our knowledge of French secondary schools and how they work, and provide a basis for new studies that will lead to further development of the set of indicators itself.

Methodological and policy choices have affected the manner in which the system is being implemented in a number of important ways: these will be discussed following a description of the broad context of the operation and the factors that have, only quite recently, made this whole project possible. Lastly, the set of indicators that should be available to all schools in France in 1995 is presented.



1. Le contexte

Il se caractérise, à la fois, par une exigence institutionnelle forte qui fait de l'élaboration et de l'évaluation des projets d'établissement une nécessité légale, par une demande sociale de plus en plus exigeante et pressante quant aux résultats et à l'efficacité des établissements scolaires et par la mise en place de premiers outils qui ont permis le développement, sous des formes multiples, d'une culture et de pratiques effectives d'évaluation dans le système éducatif français; pour autant, l'observation montre que les pratiques en matière d'auto-évaluation des établissements secondaires restent limitées.

Une nécessité légale

L'autonomie des établissements, la nécessité pour eux d'élaborer un projet, de l'évaluer, d'en rendre compte sont, à l'issue d'une évolution qui a duré plusieurs dizaines d'années, très largement acceptées aujourd'hui dans la société et le système éducatif français. La loi d'orientation sur l'éducation du 10 juillet 1989 le constate et le formalise dans la partie qu'elle consacre aux établissements d'enseignement :

« Les écoles, les collèges, les lycées d'enseignement général et technologique et les lycées professionnels élaborent un projet d'établissement. Celui-ci définit les modalités particulières de mise en œuvre des objectifs et des programmes nationaux. Il

fait l'objet d'une évaluation. Il précise les activités scolaires et périscolaires prévues à cette fin. Les membres de la communauté éducative sont associés à l'élaboration du projet qui est adopté par le conseil d'administration ou le conseil d'école (...).

L'exigence d'un projet d'établissement qui organise, dans une situation spécifique dont il faut pouvoir établir l'analyse et le diagnostic, l'ensemble des activités concourant à la mise en œuvre des objectifs et des programmes nationaux trouve ainsi sa légitimité au plus haut niveau; l'exigence d'une évaluation de ce projet est affirmée avec la même force. Quant à l'association de tous les partenaires de la communauté éducative, elle suppose, pour être effective, une transparence minimale de l'analyse, du diagnostic, de l'évaluation et du suivi.

Il ne s'agit pas de tracer l'évolution qui a progressivement conduit à cette exigence; disons simplement qu'elle accompagne la reconnaissance des diversités locales par l'institution et qu'elle aboutit à légitimer l'établissement scolaire, comme entité spécifique d'éducation ayant un effet propre, ce qui, dans le système éducatif français est relativement récent. «L'établissement est la cellule de base de l'efficacité du système éducatif» déclarait le ministre de l'Éducation nationale en juin 1994, en annonçant les mesures qu'il retenait dans le cadre du «Nouveau contrat pour l'école».

En tout état de cause, la question des outils dont les établissements ont besoin pour élaborer, évaluer et suivre leur projet se trouve ainsi posée dans un cadre conceptuel différent de celui qui avait conduit, dans la décennie 70, à élaborer et mettre en place des tableaux de bord d'établissements secondaires. Au risque d'être caricatural, on pourrait dire que ces travaux étaient inspirés par une recherche de la rentabilité des établissements, alors que les dispositifs actuels sont fondés sur la volonté d'en apprécier l'efficacité éducative globale.

Une demande sociale de plus en plus exigeante et pressante quant aux résultats et à l'efficacité des établissements

Depuis 1980, le succès que rencontre la publication dans la presse des résultats au baccalauréat «lycée par lycée» et de nombreux palmarès et classements d'établissements, plus ou moins sérieusement étayés, témoigne de cette demande.

Elle résulte certainement de la combinaison de plusieurs facteurs: tout d'abord, elle apparaît au moment où se développe un fort courant de doute sur l'école, accusée de contribuer au développement du chômage, de ne plus remplir ses missions, etc.; ensuite, elle correspond au développement du consumérisme scolaire et à la prise de conscience, déjà évoquée ci-dessus, de la diversité des situations et des performances des établissements; enfin, elle est bien évidemment liée aux mesures prises, à partir de 1986, pour augmenter sensiblement la marge dont disposent les parents pour choisir un collège ou un lycée public pour leurs enfants. Cette demande est considérée aujourd'hui comme relativement légitime et elle est, en tout cas, suffisamment forte pour que les établissements, dont un bon nombre étaient, au départ, plutôt réticents, par principe, acceptent aujourd'hui de communiquer à la presse les informations qu'elle peut souhaiter.

Comme la nécessité légale, cette demande du corps social pose à l'institution la question des outils propres à la satisfaire correctement; s'y ajoute la question de la

diffusion et de la publicité à donner aux résultats obtenus grâce à ces outils – ceci d'autant plus que le ministère de l'Éducation nationale a considérablement développé l'information et la transparence sur le fonctionnement et les résultats du système éducatif français (ministère de l'Éducation nationale, 1993, 1994a) et nourri le débat sur la façon d'apprécier l'efficacité des établissements, ce qui a fait monter le niveau d'exigence en la matière.

Les premières réalisations ont permis de développer une culture de l'évaluation au niveau des établissements secondaires

Dès 1988, la Direction de l'évaluation et de la prospective a mis en place les premières batteries d'indicateurs destinées à chaque établissement du second degré, avec un double objectif : leur restituer, sous une forme utilisable pour l'analyse et le suivi de leur situation, les données recueillies dans les enquêtes annuelles et, pour les lycées d'enseignement général et technologique, leur fournir des indicateurs d'appréciation de leurs résultats plus pertinents que les seuls taux de réussite au baccalauréat.

Ces premières batteries d'indicateurs, présentées sous forme de fiches-papier, étaient loin de constituer un ensemble correspondant à un modèle complet de l'action et de l'organisation d'un établissement scolaire : elles comprenaient, pour les lycées d'enseignement général et technologique, des éléments d'appréciation du niveau scolaire de la population accueillie (âges moyens et origines scolaires des élèves présents dans l'établissement), des données relatives aux ressources, notamment en termes d'options facultatives proposées aux élèves, et quelques indicateurs de fonctionnement (taille des classes, enseignements dispensés). Les uns et les autres étaient assortis de références nationales et académiques permettant à l'établissement de se situer.

La mise à disposition de ces indicateurs de l'ensemble des établissements a eu plusieurs mérites incontestables, notamment auprès des chefs d'établissement : elle leur a prouvé que les recueils d'informations qui leur étaient demandés par l'administration n'étaient pas seulement une charge de travail pour eux, mais pouvaient déboucher sur un service qui leur était rendu ; elle leur a permis de travailler sur des indicateurs, c'est-à-dire des informations construites, et ceci, tant de façon synchronique que diachronique ; elle leur a montré la différence entre les résultats d'un établissement et les facteurs qui y conduisent. Un certain nombre de chefs d'établissement ont complété cette première batterie d'indicateurs nationaux avec d'autres indicateurs de leur cru, pendant que des responsables académiques de l'Éducation nationale, plus avancés dans la mise en place de systèmes d'informations comprenant un repérage individuel des élèves, mettaient à la disposition des établissements de leur ressort des indicateurs fondés sur le suivi de cohortes réelles d'élèves.

De plus, une « fiche baccalauréat » proposait aux lycées d'enseignement général et technologique trois indicateurs qui permettaient une approche plurielle de leurs performances, voire de leur efficacité dans la préparation de cet examen (ministère de l'Éducation nationale, 1994c). Il s'agissait 1) du taux traditionnel de réussite à l'examen, bien sûr, mais aussi de deux indicateurs qui illustraient d'autres dimensions et dont le calcul n'était pas trop sensible à l'existence de flux d'élèves importants entre établissements, à savoir 2) parmi les élèves entrés au lycée à 15 ans (l'âge normal), la proportion de ceux

ayant obtenu le baccalauréat en trois ans (durée normale de la scolarité) dans le lycée et 3) parmi les élèves quittant l'établissement, à quelque niveau que ce soit et quelles qu'en soient les raisons, la proportion de ceux le quittant avec le baccalauréat.

Le deuxième indicateur, qui mesurait le taux de réussite dans le lycée, sans redoublement, pour des élèves arrivés à l'âge normal en seconde, portait sur une population relativement homogène d'un lycée à l'autre et permettait ainsi une approche de l'efficacité globale du lycée à l'égard de cette population. Le troisième indicateur avait pour ambition de corriger l'impression donnée par le seul taux de réussite à l'examen dans un pays où on peut penser que certains lycées sélectionnent leurs élèves à l'entrée et/ou en cours de cycle; il estimait les chances qu'un élève, entré en seconde dans un lycée, a de sortir de ce même lycée avec le baccalauréat, quelle que soit la durée nécessaire pour ce faire. Un lycée qui élimine ses élèves médiocres ou moyens aura, en effet, une valeur faible pour cet indicateur.

La qualité du système d'information n'a pas permis d'envisager la publication des indicateurs de cette « fiche baccalauréat » qui étaient simplement transmis aux établissements; les travaux qu'ils ont permis de conduire sont restés prudents et ont pris en compte, non pas la valeur absolue des indicateurs, mais, pour chacun d'eux, le classement des lycées en trois groupes. Ils ont mis à jour un point important: 13 pour cent seulement des lycées relevaient d'un classement dans le premier tiers pour les trois indicateurs, la même proportion relevait d'un classement dans le dernier tiers, toujours pour les trois indicateurs. Plus des deux tiers des lycées avaient des classements différents pour au moins deux (59 pour cent) ou trois (9 pour cent) indicateurs.

Cette « fiche baccalauréat » a ainsi permis de faire avancer la réflexion et la pratique en matière d'appréciation des performances des établissements sur deux plans: en établissant nettement la différence entre la mesure d'un résultat à l'examen et l'estimation de la performance d'un lycée, d'une part, en montrant qu'il était nécessaire de prendre plusieurs points de vue pour apprécier la réalité complexe que sont les performances d'un établissement, d'autre part. Elle a également permis aux établissements d'argumenter sur leurs points forts lorsqu'ils se trouvaient confrontés à des jugements à l'emporte-pièce fondés sur la publication d'un palmarès ou d'un classement univoque par tel ou tel organe de presse.

Le développement de pratiques d'évaluation chez les enseignants à partir d'outils standardisés

On a surtout évoqué jusqu'ici, ce qui est logique, les exigences et les premiers outils mis en place en matière d'évaluation des établissements, mais, pour brosser un tableau d'ensemble du contexte français, il est nécessaire de dire quelques mots sur les enseignants. Le développement, depuis 1989, d'évaluations de masse des élèves à quelques niveaux importants de la scolarité (début de la troisième année de l'enseignement élémentaire, entrée au collège puis, à partir de la rentrée scolaire 1992, entrée au lycée) a contribué à faire pénétrer une « culture de l'évaluation » dans les pratiques des enseignants, non pas tant pour noter ou classer les élèves, ce qui n'aurait rien eu de nouveau, que pour faire le point de leurs acquis et de leurs besoins. La fourniture aux enseignants de « banques » d'items d'évaluation, adaptées aux objectifs des différents niveaux de la

scolarité et développées par la Direction de l'évaluation et de la prospective depuis plusieurs années, va dans le même sens.

Bien qu'il reste des résistances chez certains enseignants, on a tout lieu de penser que la mise à disposition de ces outils et le développement des pratiques d'évaluation des élèves contribuent à donner une image de l'évaluation qui permet, tout autant, sinon plus, de réguler les pratiques pédagogiques plutôt que de prononcer un jugement extérieur.

L'auto-évaluation des établissements reste difficile et très minoritaire

Malgré la conjonction d'une exigence légale, d'une forte demande sociale, de la mise en place de premiers outils auprès des chefs d'établissement et du développement d'une culture de l'évaluation chez les enseignants, les différentes observations conduites pour apprécier les pratiques effectives en matière d'auto-évaluation des établissements donnent l'impression d'un manque.

Dans un rapport conjoint, les Inspections générales de l'Éducation nationale et de l'Administration de l'Éducation nationale (1991) notent que « l'existence d'une auto-évaluation constitue, en général, l'un des critères de validation des projets [par les rectorats]. Toutefois, la volonté affirmée de la mettre en place se heurte à la difficulté de trouver des indicateurs pertinents et au manque de savoir-faire des intéressés ». L'impression n'est cependant pas totalement négative puisque, dans son rapport sur l'évaluation d'une centaine de lycées, l'Inspection générale de l'Éducation nationale (1991) estime que « si la préoccupation de l'évaluation figure explicitement dans la plupart des projets [des lycées], elle demeure au stade de l'intention. Des processus d'évaluation précis, intégrés au projet, n'existent que dans 15 pour cent des cas ».

Dans une étude réalisée pour la Direction de l'évaluation et de la prospective, auprès de quelques collèges et lycées sélectionnés (parce que les responsables locaux de l'Éducation nationale estimaient que c'étaient des établissements où la volonté – sinon les pratiques – en matière d'auto-évaluation était forte), il apparaît manifestement que les plus motivés des acteurs – les chefs d'établissement – rencontrent de grandes difficultés lorsqu'ils cherchent à fédérer l'ensemble des évaluations partielles de toutes les actions conduites dans leur établissement en une auto-évaluation globale de celui-ci (Fixari et Kletz, 1993). Si les auteurs de l'étude proposent, pour expliquer ces difficultés, un certain nombre de facteurs touchant au fonctionnement du système éducatif français, ils notent également que l'absence d'outils et de cadre de référence jouent un rôle important.

2. Les orientations retenues pour la mise en place du dispositif

Elles prennent en compte des éléments de l'analyse que l'on peut faire à partir des données de contexte qui viennent d'être présentées. Pour ce faire, elles se fondent sur quelques principes méthodologiques et font appel à plusieurs conditions de mise en œuvre et de soutien de l'opération.

Se placer dans une optique d'aide et de service rendu aux établissements

On a vu plus haut qu'un des éléments du diagnostic que l'on pouvait faire de la situation actuelle était la difficulté, pour les établissements, de trouver des indicateurs pertinents et leur manque de savoir-faire ; on peut ajouter que, lorsque des établissements ont pu et su élaborer des indicateurs pertinents, il est rare que cet effort s'inscrive dans la durée, ceci pour différentes raisons : lassitude ou impossibilité de recueillir chaque année les données nécessaires, problèmes de mise en forme de ces données, ou, plus simplement, changement dans l'équipe de direction de l'établissement.

On a donc résolument pris le parti d'apporter une aide aux établissements, tout d'abord en leur fournissant une base d'indicateurs calculés à partir des informations recueillies dans les systèmes d'information et de gestion nationaux, ensuite, en assortissant ces indicateurs de références nationales et académiques qui leur permettent de se situer par rapport aux autres établissements et, enfin, en leur donnant les moyens de calculer, de suivre et de garder en mémoire les indicateurs spécifiques qu'ils souhaitent élaborer pour évaluer et réguler leur projet. Ceci se matérialise concrètement par la transmission télématique aux établissements d'indicateurs et de références calculés au niveau national et par la mise à disposition d'un logiciel de traitement et de présentation de données dont une bonne partie est directement prélevée dans les bases d'information de l'établissement ; s'ils le souhaitent, le logiciel leur permet également d'élaborer d'autres indicateurs.

La fourniture d'indicateurs et de références déjà élaborés constitue, certes, un aspect matériel important pour les établissements, mais elle a également une valeur symbolique et politique essentielle. En effet, la refonte récente du système d'information sur les élèves et la scolarité, et de pans importants du système de gestion de l'Éducation nationale, s'est traduite pour les établissements – tout particulièrement pour leurs équipes de direction et de gestion – par de lourdes tâches d'apprentissage et de mise en place de procédures nouvelles et par un recueil de données particulièrement important. La fourniture du dispositif d'indicateurs représente pour les établissements le premier résultat de ces efforts et leur donne la preuve tangible que les informations recueillies ne sont pas uniquement destinées à l'information des décideurs et des gestionnaires centraux et académiques, mais ont aussi une utilité directe pour eux.

Plus fondamentalement, ce dispositif technique correspond à la mise en œuvre de plusieurs orientations méthodologiques et politiques essentielles en ce qui concerne l'évaluation des établissements scolaires français.

Un outil qui permet à l'établissement, à la fois, de se rendre compte de son fonctionnement et d'en rendre compte

Le nom donné au système mis en place : « IPES, Indicateurs pour le pilotage des établissements secondaires » met l'accent, au premier chef, sur l'évaluation interne de l'établissement. C'est un outil destiné au chef d'établissement et à son équipe, un outil de diagnostic de la situation et du fonctionnement de l'établissement, d'évaluation de la politique qui y est conduite et de ses résultats, de régulation et de suivi de son projet.

Mais ce doit être aussi un outil de compte rendu de son activité et de ses résultats ; la notion de compte rendu doit, ici, être prise au plein sens du terme. Les collèges et les lycées sont des unités du service public d'Éducation nationale et ont, à ce titre, des « comptes à rendre » à leur tutelle, bien sûr, mais aussi, plus largement, à la collectivité.

Cette double finalité de l'évaluation des établissements – et du dispositif mis en place pour la concrétiser – s'est dégagée progressivement, au fil du temps et des travaux. C'est aujourd'hui un point fort et l'on verra plus loin que la publication par le ministère de l'Éducation nationale de trois indicateurs de performance des lycées – ce qui instaure une pratique résolument nouvelle pour le système éducatif français – vient d'être réalisée au moment même où les premiers collèges et lycées recevaient leurs indicateurs de pilotage.

Un outil qui permet une évaluation qui se fonde, à la fois sur une base d'indicateurs relativement complète et commune à l'ensemble des établissements et sur des indicateurs spécifiques à chaque établissement et à son projet

On retrouve là, sous une autre forme, la dualité évoquée ci-dessus.

On doit disposer pour chaque établissement – et chaque établissement doit disposer – d'indicateurs correspondant à ce que l'on estime indispensable de prendre en compte, sur le plan méthodologique, mais aussi politique, pour évaluer – et pour qu'il évalue lui-même – son fonctionnement, ses résultats et ses performances. On lui propose donc – et dans une certaine mesure on le contraint à utiliser – une base standard d'indicateurs qui constitue, en quelque sorte, un compromis entre tout ce que l'on estimerait souhaitable d'apprécier pour évaluer aussi complètement que possible le fonctionnement et les performances d'un établissement, et ce qu'il paraît raisonnable de faire effectivement, dans l'état actuel des informations disponibles, pour un coût tolérable (en argent, certes, mais aussi en temps de travail pour les responsables des établissements).

Cette base standard – on le verra plus loin – s'efforce évidemment de décrire, aussi exhaustivement que possible, tous les aspects d'un modèle de fonctionnement des établissements scolaires. Il faut en effet éviter certains écueils : que des comparaisons entre établissements ne prennent en compte que des aspects trop partiels de la réalité (par exemple, en ne considérant que les caractéristiques des populations qu'ils accueillent sans s'interroger sur ce qu'ils leur apportent) ou que l'on se contente d'un jugement sur les résultats sans rechercher les raisons de ceux-ci et sans envisager les infléchissements à apporter, éventuellement, à la politique de l'établissement.

Mais, au-delà de cette base standard, il importe que chaque établissement puisse décrire et évaluer des aspects particuliers de son fonctionnement ou de ses résultats, qu'il puisse se doter d'indicateurs spécifiques, à condition bien sûr, de disposer de l'information nécessaire ou de s'organiser pour la recueillir. Il est en effet normal, dans un contexte où la loi invite les établissements à « définir les modalités particulières de mise en œuvre des programmes et des objectifs nationaux », qu'on leur facilite l'évaluation et la régulation de ces modalités particulières.

Une généralisation rapide et un important soutien en formation

Le premier point correspond à un principe de mise en œuvre dont l'expérience montre qu'il est souvent décisif dans l'administration et le système éducatif français. On a donc fait le choix – pari risqué et quelque peu hasardeux, mais jusqu'à présent tenu – de mettre en place tout le dispositif, dans tous les établissements, en un an et demi. Environ 600 établissements, soit quelque 5 pour cent d'entre eux, en disposent depuis mai 1994; la totalité devraient le recevoir d'ici fin 1995.

Le second point implique un effort particulièrement important puisque l'objectif en est que tous les chefs d'établissement aient reçu une formation – ou tout au moins, une information – sur le dispositif et l'utilisation des indicateurs standard au moment où ils en disposeront. A cette fin, des « guides de l'utilisateur » des indicateurs ont été élaborés et deux cas pédagogiques, consacrés respectivement à un collège et à un lycée, ont été mis au point. Ils constituent le support de formations qui, en règle générale, durent deux jours et permettent aux chefs d'établissement – et dans certains cas, à un de leurs collaborateurs – de s'approprier les indicateurs, d'en connaître les règles d'élaboration et, surtout, de les utiliser en situation.

3. La batterie d'indicateurs standard

On prendra l'exemple des lycées d'enseignement général et technologique, étant entendu que chaque type d'établissement du second degré (les collèges, les lycées professionnels et les lycées d'enseignement général et technologique) dispose d'une batterie qui, à partir d'un tronc commun, prend en compte ce qui est spécifique à chaque type d'établissement.

Ces batteries d'indicateurs ont été construites en faisant très largement appel à l'expérience et aux besoins des praticiens, tout en utilisant comme toile de fond les travaux de recherche sur l'efficacité de l'école.

Dans de nombreux cas, on a dû faire largement appel à ce que l'on pourrait appeler « la règle du compromis » entre le désir que l'on pouvait avoir de construire un indicateur qui rende compte de façon aussi précise et complète que possible d'un phénomène et le souci d'utiliser au mieux les dispositifs d'informations et de gestion existants et de ne pas multiplier les saisies d'informations nouvelles, pour lesquelles les établissements sont – à juste titre – particulièrement réticents. Ensuite, on a tenu à se limiter à une vingtaine d'indicateurs, ce qui a parfois impliqué de résister à la tentation de tout apprécier, pour se limiter à quelques grands axes essentiels que les établissements pourront, comme on l'a vu plus haut, compléter en fonction de leurs besoins et de leurs priorités spécifiques. Enfin, ces indicateurs ont fait l'objet de tests dans un certain nombre d'établissements.

On s'est efforcé, dans cette tâche, de respecter deux principes : d'une part, donner un ensemble d'indicateurs qui s'intègrent dans une représentation formalisée, dans un modèle de l'action et de l'organisation des établissements scolaires et, d'autre part, présenter des indicateurs qui rendent compte, au-delà des résultats d'un établissement, de sa performance effective et de sa « valeur ajoutée ».

Décrire, aussi complètement que possible, tous les aspects d'un modèle de fonctionnement des établissements scolaires

Le modèle, tel qu'il est présenté aux chefs d'établissement, est tout à fait classique ; on peut, schématiquement, le représenter comme dans le graphique 10.1. Un établissement est, en effet, une organisation qui, dans un environnement donné, et par un certain type de fonctionnement, mobilise des ressources pour que les élèves qu'il accueille et qui sont dotés d'un certain nombre de caractéristiques, en sortent avec des caractéristiques de niveau supérieur.

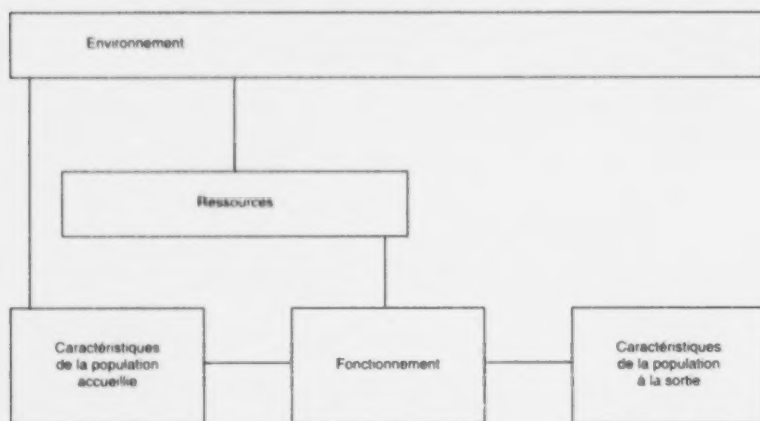
Si la batterie d'indicateurs proposée ne mesure pas l'ensemble des éléments de ce modèle (elle ne comprend pas, par exemple, d'indicateur rendant compte du « climat » de l'établissement qui, on le sait, est un élément important de son efficacité), elle comprend cependant au moins un indicateur pour chacun de ses aspects (voir tableau 10.1).

C'est par les indicateurs de résultats et surtout par la question de l'appréciation de la « valeur ajoutée » des lycées que débutera cette présentation.

Dépasser l'appréciation des résultats des établissements pour évaluer leur « valeur ajoutée »

Cette question est d'autant plus importante que, comme on l'a indiqué en début de cet article, les résultats des lycées au baccalauréat, font, chaque année en France, l'objet

Graphique 10.1. La mise en place d'un dispositif d'indicateurs pour le pilotage des établissements secondaires français
Steering indicators for French secondary schools



Source : Auteur.

Tableau/Table 10.1.
Batterie d'indicateurs standard des lycées
Standard indicators for upper secondary schools

Indicateurs de résultats	
R1 –	Taux de réussite au baccalauréat par série (brut, attendu, indice de comparaison), par âge (brut)
R2 –	Taux d'accès au niveau supérieur et au baccalauréat (taux brut, attendu, indice de comparaison)
R3 –	Proportion de bacheliers parmi les sortants du lycée
R4 –	Destination des élèves à leur sortie du lycée avec (R 4.1) et sans diplôme (R 4.2)
Indicateurs relatifs à la population prise en charge	
P1 –	Caractéristiques de l'ensemble des élèves de l'établissement à la rentrée
P2 –	Caractéristiques des élèves présents dans les classes de lycée à la rentrée : sexe (P 2.1), catégories socioprofessionnelles (P 2.2), âge moyen et proportion d'élèves en retard (P 2.3), proportion de redoublants (P 2.4); effectifs présents en première par section : âge moyen, élèves en retard (P 2.5); effectifs présents en terminale par section : âge moyen, élèves en retard (P 2.6); proportion d'élèves majeurs avant la fin de l'année scolaire (P 2.7)
P3 –	Caractéristiques des élèves entrant dans le lycée à la rentrée : sexe, âge, proportion de redoublants (P 3.1); originaires de l'enseignement privé (P 3.2)
P4 –	Origines scolaires des élèves entrant dans les sections à la rentrée
Indicateurs de ressources et de moyens	
M1 –	Moyens horaires
M2 –	Caractéristiques des enseignants
M3 –	Effectifs des élèves par option
Indicateurs de fonctionnement et d'environnement	
F1 –	Devenir des élèves en fin de seconde
F2 –	Proportion d'heures d'enseignement effectivement assurées à un élève
F3 –	Taille moyenne des divisions, comparaison établissement-académie
F4 –	Importance des services liés à l'enseignement
F5 –	Importance des services liés à la vie scolaire
F6 –	Mouvements du corps enseignant
F7 –	Accidents dans l'établissement
F8 –	Statut des lycéens
E1 –	Intensité des relations avec l'environnement économique

de publications dans la presse, et ceci généralement sous forme de palmarès ou de classements censés rendre compte des mérites et des qualités comparés des établissements.

S'il y a plusieurs années que la Direction de l'évaluation et de la prospective du ministère de l'Éducation nationale travaille à la mise au point d'indicateurs de performance des établissements, ce n'est qu'à partir de 1993 que se trouvent réunies – grâce aux améliorations importantes apportées au système d'information du ministère – les conditions jugées indispensables, du point de vue d'un service public national, pour

envisager, à l'échelle de l'ensemble des établissements secondaires, une solution satisfaisante qui puisse, selon un des principes énoncés plus haut, à la fois être utile à l'auto-évaluation des établissements et donner lieu à une large diffusion publique.

Tout d'abord, il est possible de présenter, dans des conditions correctes de fiabilité, plusieurs indicateurs pour chaque lycée, et ceci en insistant sur le fait qu'ils doivent être examinés concurremment, parce qu'ils ont chacun un intérêt particulier et mettent l'accent sur des aspects différents – et complémentaires – des performances des établissements. Une vision trop partielle serait aussi partielle et guère plus fondée que les rumeurs qui peuvent circuler à leur sujet; ensuite, il est également possible de proposer et calculer, parmi cet ensemble, des indicateurs qui donnent une appréciation de l'efficacité d'un lycée, c'est-à-dire de ce qu'il apporte à ses élèves (et de ce que ceux-ci tirent de cet apport) en éliminant, autant que possible, l'incidence de facteurs sur lesquels l'efficacité de l'établissement ne saurait jouer comme, par exemple, le plus ou moins grand retard scolaire que peuvent avoir les élèves lorsqu'il les accueille.

Pour illustrer cette démarche et ses résultats, on reprendra ci-dessous les grandes lignes de la présentation des «trois indicateurs de performance, lycée par lycée» publiés par le ministère de l'Éducation nationale (1994c), à partir des résultats de la session 1993 du baccalauréat :

«Que mesurer, quels critères retenir pour apprécier les résultats d'un établissement? Il n'y a pas de réponse unique à ces questions et l'on sait que les objectifs de chaque lycéen ou de ses parents sont différents : certains privilégieront le fait d'obtenir telle série du bac et seront disposés, pour ce faire, à accepter un redoublement ou à changer d'établissement; d'autres souhaiteront effectuer toute leur scolarité lycéenne dans le même établissement, proche de chez eux; d'autres, encore, désireront obtenir un baccalauréat le plus rapidement possible, etc. C'est pourquoi, un palmarès ou un classement n'a guère de sens – en tout cas pas de sens général – et l'on peut imaginer, à la limite, une multiplicité d'indicateurs rendant compte de ce que recherchent les uns ou les autres.»

On a donc retenu plusieurs indicateurs qui proposent des approches différentes et complémentaires des performances des lycées : 1) le taux de réussite au baccalauréat, 2) le taux d'accès de seconde au baccalauréat et 3) la proportion de bacheliers parmi les sortants du lycée.

Le taux de réussite au baccalauréat est l'indicateur le plus traditionnel, le plus connu du grand public – et surtout le plus facile à établir – qui montre le nombre d'élèves du lycée reçus au baccalauréat en 1993 par rapport au nombre d'élèves présents à la première épreuve.

Ce taux sera donc flatteur pour un établissement qui sélectionne parmi les élèves qui veulent entrer en seconde, puis parmi les élèves de seconde et de première, ceux qu'il estime aptes à obtenir le baccalauréat et qui laisse d'autres lycées accueillir ceux dont il estime faibles les chances de succès. En revanche, le taux de réussite d'un établissement dont les élèves de terminale sont nombreux à obtenir le baccalauréat – non pas toujours du premier coup, mais parfois après un, voire plusieurs, redoublements – sera moins favorable, alors qu'une proportion importante de ces élèves seront bacheliers lorsqu'ils quitteront le lycée.

Si cet indicateur est donc tout à fait insuffisant pour rendre compte de l'efficacité d'ensemble d'un lycée, il ne saurait être négligé et il a l'avantage de pouvoir être établi aisément et de pouvoir être présenté pour chaque série du baccalauréat.

Le taux d'accès de seconde au baccalauréat est un indicateur nouveau (qui peut, aujourd'hui, être établi pour tous les établissements publics grâce au nouveau système d'information portant sur les élèves du second degré). Il évalue, pour un élève de seconde, la probabilité qu'il obtienne le baccalauréat à l'issue d'une scolarité entièrement effectuée dans le lycée, et ceci quel que soit le nombre d'années nécessaire. Il n'est pas établi à partir d'une cohorte réelle d'élèves, mais résulte d'un calcul effectué sur une cohorte instantanée (cf. encadré 10.1), ce qui permet de rendre compte de la politique conduite par le lycée pour l'année où il est établi et non pas de la résultante des politiques conduites au cours des trois, quatre, voire cinq années nécessaires à une cohorte d'élèves pour parcourir tout le cycle de la scolarité lycéenne.

Contrairement au précédent, cet indicateur s'intéresse donc à l'ensemble de la scolarité dans le lycée. Il ne compte au crédit de celui-ci que les élèves qui sont susceptibles d'y effectuer une scolarité complète couronnée de succès, même si cette scolarité comprend un, voire plusieurs redoublements; en revanche, les élèves qui ont dû quitter le lycée, soit parce que celui-ci n'offrait pas la section qu'ils souhaitaient préparer, soit parce qu'on les a orientés vers une telle section pour les éliminer, ou, tout bonnement, parce qu'ils ont démenagé, sont considérés, du point de vue de l'indicateur, comme des élèves que le lycée n'a pas su (ou pas pu) conduire au baccalauréat, même si, en fin de compte, ils ont obtenu ce diplôme dans un autre établissement.

On doit donc garder à l'esprit que si cet indicateur est beaucoup plus pertinent que le taux de réussite au baccalauréat pour apprécier l'efficacité globale d'un lycée, il peut être influencé par des facteurs sur lesquels le lycée n'a que peu ou pas de prise et qui ne relèvent donc pas, à proprement parler, de son efficacité; ainsi, un lycée qui propose toutes les sections du baccalauréat et qui est relativement isolé a plus d'atouts pour garder ses élèves jusqu'en terminale (et jusqu'à l'obtention d'un baccalauréat) qu'un lycée que de nombreux élèves quittent en cours de cycle parce qu'il ne leur propose pas la section qu'ils souhaitent ou parce qu'il est implanté dans une zone où la concurrence d'autres établissements est très forte.

La proportion de bacheliers parmi les sortants présente, parmi les élèves qui ont quitté l'établissement une année donnée, quelles qu'en soient les raisons, la proportion de ceux qui l'ont quitté avec le baccalauréat: lorsque cet indicateur porte sur les élèves qui quittent le lycée à tous les niveaux (seconde, première et terminale), il rend compte, comme le précédent, de l'ensemble de la scolarité dans l'établissement, mais d'une façon différente, puisqu'il évalue les chances qu'un élève de seconde a de quitter le lycée en étant titulaire du baccalauréat, et ceci quel que soit le nombre d'années qu'il y passe; lorsqu'il ne porte que sur les élèves qui quittent la terminale, il estime les chances qu'un élève inscrit en terminale a de quitter le lycée en étant bachelier, y compris en redoublant la terminale dans l'établissement. Un lycée qui sélectionne ses élèves et qui n'autorise pas (ou très peu) le redoublement aura, pour cet indicateur, une valeur plus faible qu'un lycée qui laisse plusieurs chances à ses élèves.

Encadré 10.1

Le taux d'accès de seconde au baccalauréat
Completed upper secondary rates

Le taux d'accès de seconde au baccalauréat pour un lycée est le produit des taux d'accès de seconde en première, de première en terminale et de terminale au baccalauréat calculés pour cet établissement et pour une année donnée.

Le calcul de ces taux n'est pas fondé sur l'observation d'une cohorte réelle d'élèves entrés en seconde, mais sur ce qu'il est convenu d'appeler une cohorte instantanée. Ces taux décrivent mieux, à un moment donné, ce qui se passe dans l'établissement que ceux fondés sur des cohortes réelles. En effet, dans ce dernier cas, on est contraint d'attendre que la cohorte ait quitté le lycée (4 ou 5 ans au moins).

Le taux d'accès d'un niveau à l'autre est le rapport de la population des élèves d'un établissement quittant le niveau de départ sur la population de ceux qui accèdent au niveau supérieur dans le même établissement.

La formule est la suivante, pour l'année 1992/93 par exemple :

$$\text{Taux d'accès d'un niveau à l'autre} = \frac{\text{ESL}}{\text{EI} - \text{ERL}} \times 100$$

Avec :

EI = élèves inscrits dans le niveau de départ en septembre de l'année 1992

ERL = élèves de l'établissement qui redoublent le niveau de départ dans le lycée en septembre 1993

Pour le calcul des taux d'accès de seconde en première et de première en terminale :

ESL = élèves de l'établissement qui passent dans le niveau supérieur dans le lycée en septembre 1993

Pour le calcul du taux d'accès de terminale au baccalauréat :

ESL = élèves de l'établissement ayant obtenu le baccalauréat en juin ou en septembre 1993

Les trois taux présentés ci-dessus proposent donc des approches complémentaires de la performance des lycées, mais ils constituent des données brutes qui ne renseignent pas sur la question fondamentale : comment évaluer l'action propre du lycée, ce qu'il a « ajouté » au niveau initial des élèves qu'il a reçus. En d'autres termes, si un lycée présente une valeur élevée pour un indicateur, est-ce dû au fait qu'il a reçu de bons élèves, dotés de bonnes méthodes de travail, qui ont pu obtenir le baccalauréat sans effort particulier de sa part, ou bien, est-ce dû au fait qu'il a su, tout au long d'une scolarité, développer chez des élèves peut-être moins bien dotés au départ, les connaissances et les capacités qui ont permis leur succès?

Pour apprécier cette valeur ajoutée, il faut donc s'efforcer d'éliminer l'incidence des facteurs de réussite scolaire extérieurs au lycée, pour ne conserver que ce qui est dû à son action propre. On a retenu les deux facteurs extérieurs principaux – l'âge et l'origine socioprofessionnelle des élèves – ce qui permet de comparer la donnée brute (la valeur de l'un des taux calculés pour le lycée) et une donnée « attendue » : la valeur que prendrait ce taux si, en moyenne, tous les élèves du lycée connaissaient la même scolarité et le même succès au baccalauréat que l'élève moyen comparable – c'est-à-dire de mêmes âge et origine socioprofessionnelle – de l'ensemble des lycées (cf. encadré 10.2).

Encadré 10.2

Les taux attendus

Anticipated rates

Si le taux de réussite brut au baccalauréat se calcule simplement – il s'agit du rapport entre le nombre de bacheliers et le nombre de candidats présents à la première épreuve – le calcul du taux de réussite attendu est, quant à lui, plus complexe.

Taux de réussite attendu

Pour le calculer, on regroupe les élèves du lycée qui se sont présentés au baccalauréat en neuf sous-populations obtenues en combinant trois catégories d'âge (i) et trois regroupements d'origine sociale (j). Le taux de réussite attendu au baccalauréat est obtenu en appliquant aux effectifs de chacune des neuf sous-populations ($n_{i,j}$), les taux de réussite bruts ($t_{i,j}$) de l'ensemble des élèves de même âge et de même catégorie socioprofessionnelle constatés sur la population de référence (tous les élèves scolarisés dans les établissements publics ou privés sous contrat de l'académie – base académique – ou de France – base France).

$$\text{Taux de réussite attendu} = \frac{\sum_i \sum_j n_{i,j} \times t_{i,j}}{\sum_i \sum_j n_{i,j}}$$

Les regroupements par âge (i) correspondent aux élèves :

- sans retard scolaire,
- avec un an de retard scolaire,
- avec plus d'un an de retard scolaire,

c'est-à-dire aux élèves ayant 18 ans et moins, 19 ans ou 20 ans et plus l'année du baccalauréat.

Les regroupements de professions et catégories sociales (j) se font en trois catégories élaborées en tenant compte des résultats des études portant sur la réussite scolaire des élèves en fonction de leurs origines sociales :

(voir page suivante)

(suite)

- la catégorie dite « favorisée » comprenant les cadres et professions intellectuelles supérieures, les professions intermédiaires, cette catégorie comprend tous les enseignants;
- la catégorie dite « moyenne » comprenant les agriculteurs exploitants, les artisans, commerçants et chefs d'entreprise, les employés;
- la catégorie dite « défavorisée » comprenant les ouvriers et les inactifs.

Taux d'accès attendu

Pour calculer le taux d'accès attendu d'un niveau à l'autre, on regroupe les élèves du niveau de départ en neuf sous-populations obtenues en combinant les trois mêmes catégories d'âge (i) et les trois mêmes regroupements d'origine sociale (j) que ceux utilisés pour le calcul des taux de réussite attendus. Puis, on applique aux effectifs des neuf sous-populations du niveau de départ ($n_{i,j}$), les taux d'accès bruts ($ta_{i,j}$) de l'ensemble des élèves de même âge et de même catégorie socioprofessionnelle calculés sur la population de référence (tous les élèves d'un même niveau, quel que soit l'établissement public où ils sont scolarisés, comptabilisés au niveau académique - base académique - ou national - base France).

La formule est la suivante :

$$\text{Taux d'accès attendu d'un niveau à l'autre} = \frac{\sum_i \sum_j n_{i,j} \times ta_{i,j}}{\sum_i \sum_j n_{i,j}}$$

Le croisement de l'âge et de la catégorie socioprofessionnelle donne, en effet, une bonne approximation des chances (au sens statistique du terme) d'accès et de réussite au baccalauréat d'un élève, chances qui sont sensiblement différentes, en moyenne, selon que l'élève présente l'une ou l'autre de ces caractéristiques, comme le montre le tableau 10.2. Alors que, pour l'ensemble des élèves fréquentant les lycées d'enseignement général et technique publics et privés sous contrat, le taux moyen de réussite au baccalauréat s'établit à 73.3 pour cent à la session 1993, ce même taux de réussite est relativement contrasté si l'on tient compte de l'âge et de la catégorie socioprofessionnelle des élèves.

Le taux de réussite attendu d'un lycée au baccalauréat est donc le taux de réussite qui serait le sien si ses élèves connaissaient le même succès au baccalauréat que l'ensemble des candidats de même âge et de même origine socioprofessionnelle de tous les lycées. On tient le même raisonnement pour calculer un taux d'accès attendu de seconde au baccalauréat pour chaque lycée.

Si l'écart entre le taux brut et le taux attendu est positif, on a tout lieu de penser que le lycée a apporté aux élèves qu'il a accueillis plus que ce que ceux-ci auraient reçu s'ils avaient fréquenté un établissement moyen, ce qui est l'indice d'une bonne efficacité. Si l'écart est négatif, on aura la présomption inverse.

Tableau/Table 10.2.
**Taux de réussite nationaux selon l'âge et l'origine socioprofessionnelle
des élèves des établissements publics et privés sous contrat**
France métropolitaine et départements d'outre-mer
*National pass rates by age and socio-economic family background
of students in public and private schools*
Baccalauréat général et technologique - Session 1993 (%)

	Âge au 31.12.1993		
	18 ans et +	19 ans	20 ans et +
Catégorie socioprofessionnelle des parents			
Favorisée	86.5	72.1	67.3
Moyenne	82.0	69.1	65.7
Défavorisée	76.8	65.7	62.3
Total	83.6	69.6	65.3

Note : Alors que, pour l'ensemble des élèves fréquentant les lycées d'enseignement général et technique publics et privés sous contrat, le taux moyen de réussite au baccalauréat s'établit à 73.3 pour cent à la session 1993, ce même taux de réussite est relativement contrasté si l'on tient compte de l'âge et de la catégorie socioprofessionnelle des élèves.

Deux taux de réussite attendus et deux taux d'accès attendus sont présentés pour chaque établissement. L'un est calculé en prenant comme base l'ensemble des élèves de l'académie où est situé l'établissement (c'est le taux attendu sur une base académique), l'autre en prenant comme base l'ensemble des élèves de France (c'est le taux attendu sur une base nationale).

Il est en effet possible de tenir, pour l'ensemble des établissements de chaque académie, le même raisonnement que celui présenté ci-dessus, pour chaque lycée : dans chaque académie, les établissements n'ont, en moyenne, pas la même « valeur ajoutée » lorsque l'on compare leurs performances brutes à celles que l'on pourrait attendre d'eux si, dans chaque établissement, leurs élèves avaient connu les mêmes résultats que l'ensemble des élèves de même âge et de même origine socioprofessionnelle de toute la France, comme le montre le tableau 10.3.

On constate aisément à la lecture du tableau 10.3 que certaines académies présentent des résultats d'apparence flatteuse, mais que ceci est essentiellement dû aux caractéristiques de la population scolaire qu'accueillent leurs établissements, alors que d'autres sont dans une situation inverse. On constate également que, en moyenne, les établissements de nombreuses académies apportent aux élèves qu'ils ont présentés au baccalauréat une « valeur ajoutée » positive (leur taux de réussite brut est supérieur à leur taux de réussite attendu), alors que, si l'on s'intéresse à leurs performances en matière d'accès de seconde au baccalauréat, il n'en est plus de même (leur taux d'accès brut est inférieur à leur taux d'accès attendu). En fait, les situations sont diverses et contrastées et ce qui est vrai pour l'ensemble des établissements d'une académie l'est tout autant, sinon plus, pour chaque

Tableau/Table 10.3.

**Taux de réussite et taux d'accès bruts et attendus des établissements publics
et privés sous contrat de chaque académie¹**

France métropolitaine et départements d'outre-mer

*"Baccalauréat" pass rates and completed upper secondary rates,
actual and anticipated, in public and private schools, by school area¹*

Metropolitan France and overseas departments

Académies	Taux de réussite au baccalauréat		Taux d'accès de seconde au baccalauréat ²	
	Taux brut (%)	Taux attendu (%)	Taux brut (%)	Taux attendu (%)
Aix-Marseille	70.3	73.6	61.1	64.9
Amiens	69.2	72.9	—	—
Besançon	71.4	73.0	67.7	64.6
Bordeaux	75.2	73.6	66.0	65.3
Caen	72.5	73.0	—	—
Clermont-Ferrand	75.7	73.5	68.2	66.1
Corse	71.8	75.1	—	—
Créteil	67.6	73.2	—	—
Dijon	73.7	73.4	64.6	66.3
Grenoble	74.8	73.8	64.8	66.6
Lille	67.4	73.4	60.6	65.4
Limoges	69.5	73.3	—	—
Lyon	74.9	74.0	63.7	65.6
Montpellier	74.9	74.2	66.6	65.6
Nancy-Metz	69.5	72.7	66.3	65.6
Nantes	76.6	73.7	69.1	65.9
Nice	73.0	74.6	59.6	66.8
Orléans-Tours	74.6	73.3	66.6	66.3
Paris	79.1	75.3	—	—
Poitiers	75.6	73.2	—	—
Reims	68.9	72.8	—	—
Rennes	77.9	72.9	69.5	64.2
Rouen	69.2	72.2	61.8	62.9
Strasbourg	76.2	75.0	65.8	68.4
Toulouse	79.9	74.0	70.6	66.7
Versailles	78.4	74.2	66.1	65.4
France métropolitaine	73.6	73.6	65.3	65.3
Antilles-Guyane	66.9	71.8	—	—
La Réunion	57.9	73.0	—	—
France métropolitaine + DOM	73.3	73.3	—	—

1. Ces taux correspondent à des taux qui reflètent la situation de l'établissement moyen de l'académie. S'agissant des taux de réussite, ils ne considèrent que les élèves des établissements publics et privés sous contrat. En outre, en ce qui concerne les taux d'accès, ils ne prennent en compte que les élèves qui ont effectué leur scolarité du début du cycle secondaire jusqu'au baccalauréat dans le même établissement public; ils sont donc différents des taux publiés ailleurs qui sont établis en prenant en compte l'ensemble des élèves qui ont effectué leur scolarité jusqu'à l'obtention du baccalauréat dans la même académie.

2. Pour la session 1993, ces taux ne peuvent être calculés que pour 18 académies.

lycée, ce qui confirme la vanité des classements ou des palmarès qu'il est si tentant d'établir sur un critère unique, ainsi que la nécessité de multiplier les points de vue sur la délicate question de l'appréciation des résultats – et *a fortiori* – des performances des établissements. Ceci est d'autant plus vrai que l'on connaît encore mal la stabilité dans le temps de ces indicateurs.

Ainsi, dans la batterie d'indicateurs standard, l'indicateur pertinent en matière de réussite au baccalauréat (R 1) est, non pas le taux brut de réussite, mais l'écart entre ce taux brut et les taux de réussite attendus. Un écart positif dans une série donnée permet de penser que les terminales de cette série ont une bonne efficacité, tout au moins si tous les élèves se sont effectivement présentés au baccalauréat. Bien entendu, cet écart peut ne pas être de même signe selon qu'il s'agit du taux attendu sur une base académique ou du taux attendu sur une base nationale.

De même, l'indicateur pertinent en matière d'accès de seconde au baccalauréat (R 2) est l'écart entre le taux d'accès brut et les taux d'accès attendus. Un écart positif permet de penser que le lycée a une bonne efficacité globale, puisqu'il conduit au baccalauréat plus d'élèves de seconde qu'on ne s'y attendrait.

Quant à la proportion de bacheliers parmi les élèves sortant du lycée (R 3), elle n'est pas confrontée à des taux attendus, mais à deux moyennes : cette même proportion pour l'ensemble des lycées de l'académie, d'une part, et pour toute la France, d'autre part. On ne saurait tirer de cette seule confrontation avec des moyennes des conclusions sur l'efficacité de l'établissement.

Si la publication d'indicateurs de performance des lycées se limite aux trois indicateurs ci-dessus, la batterie standard fournie aux lycées en propose un quatrième (R 4) qui présente un aperçu du devenir des élèves lorsqu'ils quittent le lycée (en termes d'études poursuivies ou autres : stage, recherche d'emploi, inactivité, service militaire) et ceci indépendamment du moment où ils quittent l'établissement et qu'ils y aient ou non obtenu le baccalauréat.

Par rapport aux trois indicateurs précédents qui, tous, envisageaient comme critère principal d'appréciation des résultats du lycée, l'obtention ou non du baccalauréat par les élèves qui le fréquentent, celui-ci propose une vision un peu plus longue puisqu'il renseigne, d'une part, sur les voies auxquelles les bacheliers issus du lycée ont pu accéder et, d'autre part, donne une idée du devenir – scolaire ou autre – des élèves qui l'ont quitté en cours de cycle ou sans avoir obtenu le baccalauréat à l'issue d'une scolarité complète.

On s'efforce ainsi de donner une vision relativement complète, et généralement nuancée, des résultats et de l'efficacité d'un établissement; mais ce qui importe, en définitive, pour les équipes de direction et les enseignants des lycées, ce n'est pas de s'en tenir au seul jugement sur les performances de leur établissement, mais de disposer d'outils qui puissent les aider à en analyser les causes et à envisager les modifications et les actions propres à les améliorer; en bref de disposer de moyens supplémentaires pour élaborer, mettre en œuvre et suivre leur projet d'établissement. C'est en considérant ces objectifs que le dispositif qui leur est proposé prend tout son sens et que l'on s'est efforcé de faire en sorte qu'il puisse rendre compte, aussi complètement que possible, de leur fonctionnement.

***Donner aux établissements des moyens de comprendre leur fonctionnement,
d'apprécier leurs contraintes et leurs ressources***

Parmi les indicateurs proposés dans la batterie standard, certains appellent peu de commentaires sur le plan méthodologique : c'est notamment le cas des indicateurs relatifs à la population accueillie dans l'établissement (P 1, P 2, P 3, P 4) qui donnent à celui-ci une image des principales caractéristiques socio-démographiques et scolaires de cette population. Des indicateurs de ce type sont fournis depuis plusieurs années aux établissements, mais les progrès du système d'information sur la scolarité des élèves permettent, depuis 1993, de restituer à tous les établissements publics des données beaucoup plus pertinentes qu'auparavant, qui portent notamment sur les flux d'élèves accueillis aux différents niveaux de la scolarité et enrichissent considérablement l'appréciation qu'ils ont du niveau de la population qui leur est confiée. En particulier, le fait que les établissements disposent automatiquement de données sur les élèves entrant dans leurs classes constitue un élément nouveau particulièrement apprécié : ces entrants sont-ils plutôt des redoublants, des élèves réorientés ou des élèves issus de la classe précédente? Viennent-ils d'établissements publics ou privés? Le fait que ces indicateurs soient assortis de leur valeur moyenne au niveau national et académique permet aussi aux lycées de se situer par rapport à des références significatives.

S'ils proposent différentes approches des caractéristiques de la population scolaire du lycée, aucun de ces indicateurs ne rend directement compte du niveau de connaissances acquis par les élèves lorsqu'ils le rejoignent; chaque lycée peut cependant construire aisément un tel indicateur et suivre son évolution dans le temps en utilisant les résultats des tests d'évaluation à l'entrée en seconde de ses élèves, tests qui lui sont proposés à chaque rentrée depuis 1992. Mais les résultats de ces tests ne font pas – tout au moins pour l'instant – l'objet d'une collecte au niveau national ou académique et il n'est donc pas possible de fournir aux établissements des références qui leur permettent de situer le niveau et le degré d'hétérogénéité de leurs élèves par rapport à ceux des autres lycées. En revanche, une telle collecte est organisée (sur échantillon) pour les évaluations pratiquées à l'entrée des collèges; ceux-ci disposeront donc, au titre de la batterie d'indicateurs standards, d'un indicateur qui leur permettra d'apprécier le niveau moyen et le degré d'hétérogénéité des acquis de leurs élèves à l'entrée en sixième, en français et en mathématiques, relativement au niveau moyen et au degré d'hétérogénéité des élèves de l'ensemble des collèges.

En ce qui concerne les ressources dont disposent les établissements, les indicateurs proposés se limitent à quelques points que les établissements ne peuvent aborder par leurs propres moyens. Ainsi, aucun indicateur financier ne figure dans la batterie standard (les établissements en fabriquent aisément eux-mêmes et disposent fréquemment de données fournies par les rectorats à ce sujet). Mais les établissements disposeront d'un indicateur présentant les principales caractéristiques structurelles et socio-démographiques de leurs enseignants (M 2) et d'un autre leur permettant d'apprécier l'importance de la « palette » des options qu'ils offrent aux élèves de seconde et de première et la façon dont les élèves réagissent à cette offre en se répartissant entre ces options (M 3). Pour ces indicateurs également, la fourniture de références aux niveaux national et académique permet au lycée de se situer par rapport aux établissements voisins ou comparables.

Un indicateur de ressources : les moyens horaires des établissements

Il est, en outre, proposé aux établissements un indicateur relatif aux moyens humains dont ils disposent (M 1), et ceci non seulement en heures d'enseignement par élève, ce qui n'est pas nouveau, mais en termes de présence d'adultes par élève (enseignement, éducation, surveillance, etc.) ce qui, dans le contexte français, discrimine beaucoup plus les établissements que les heures d'enseignement par élève qui font l'objet d'un encadrement strict et d'une répartition relativement égalitaire (cf. tableau 10.4).

Table/Tableau 10.4.

Indicateur relatif aux moyens horaires (M 1)
Références nationales pour 1992-93 issues d'une enquête
sur un échantillon national d'établissements

Indicator M 1

National reference data for 1992-93, from a survey on a nationwide sample of schools

Type d'établissement	Lycées	Lycées professionnels	Collèges
Heures hebdomadaires d'enseignement par élève	1.3	2.2	1.1
Heures hebdomadaires d'enseignement et d'éducation par élève	1.8	3.0	1.9

Pour ce faire, on considère d'abord les heures d'enseignement inscrites à l'emploi du temps hebdomadaire des classes, c'est-à-dire le nombre d'heures passées devant les élèves par les professeurs au titre de leur service hebdomadaire, ce qui permet d'établir le nombre d'« heures hebdomadaires d'enseignement par élève » de l'établissement qui correspond à un ratio couramment utilisé en France pour répartir les moyens en enseignants entre les établissements.

Puis, pour établir le nombre d'heures hebdomadaires d'enseignement et d'éducation dont bénéficie en moyenne chaque élève, on ajoute :

- d'abord, les heures que certains professeurs assurent face aux élèves en sus de leur service hebdomadaire, quelle que soit la nature de leur activité (enseignement, animation péri-éducative, etc.) pour peu qu'ils en soient rémunérés par le ministère de l'Éducation nationale ou par tout autre organisme payeur (collectivité territoriale, par exemple);
- ensuite, les heures assurées par les personnels de direction, les personnels d'éducation et de surveillance, les documentalistes, les assistants de langue étrangère, etc.;
- enfin, les heures effectuées par d'autres personnes, quelles qu'elles soient, qui ont une action pédagogique ou éducative en présence des élèves (par exemple, des militaires du contingent peuvent effectuer leur service national en assurant, dans

certains établissements scolaires considérés comme « sensibles » des tâches de surveillance, d'animation, d'aide aux devoirs, etc.).

Dans l'état actuel des choses, cet indicateur ne tient pas compte du rôle éducatif que peuvent assurer les personnels d'administration, techniciens, ouvriers, de service et de santé (ATOSS) présents dans les établissements scolaires. Il le fera dès que les bases de données relatives à ces personnels le permettront.

La proportion d'heures d'enseignement dispensées à un élève

C'est dans le domaine du fonctionnement des établissements que sont proposés les indicateurs les plus nouveaux parce qu'ils s'efforcent de donner aux établissements des données aussi objectives que possible dans des domaines peu abordés jusqu'ici et/ou parce qu'ils veulent rendre compte de phénomènes qualitatifs. C'est en particulier le cas de l'indicateur relatif à la « proportion d'heures d'enseignement effectivement assurées à un élève » (F 2) et de ceux qui rendent compte des services liés à l'enseignement (F 4) et à la vie scolaire (F 5), du statut des lycéens (F 8) et de l'intensité des relations avec l'environnement économique (E 1).

Le premier de ces indicateurs permet d'apprécier la proportion d'heures d'enseignement perdues par l'élève moyen de l'établissement, relativement au nombre d'heures d'enseignement qu'il aurait dû recevoir au cours de l'année scolaire, si rien n'était venu perturber le déroulement de l'emploi du temps tel qu'il avait été prévu au début de l'année.

Selon leurs origines, les pertes d'heures appellent un constat et des mesures différenciées; l'indicateur en distingue donc quatre catégories :

- celles dues au fonctionnement du système éducatif : organisation et déroulement des examens, des procédures d'orientation, entraînant ou non la fermeture totale de l'établissement, nominations tardives de professeurs, voire fermeture de l'établissement par suite d'une décision administrative ou d'une situation de force majeure, etc.;
- celles liées au perfectionnement professionnel des enseignants : absences non remplacées essentiellement pour participation à des actions de formation;
- celles liées au non-remplacement d'enseignants absents pour des raisons individuelles (maladie, maternité, etc.);
- enfin, celles liées aux absences des élèves eux-mêmes.

Dans la mesure où la détermination précise du nombre d'heures dues à l'élève moyen du lycée est un exercice difficile et dont le résultat pourrait être quelque peu incertain, on a renoncé à l'établir en valeur absolue. L'indicateur n'est donc fourni qu'en valeur relative.

Il présente pour les établissements un grand intérêt dans la mesure où il permet, à propos d'une question délicate sur laquelle les responsables d'établissement et enseignants sont souvent questionnés, voire interpellés, de tenir un débat correctement éclairé et d'éviter les jugements stéréotypés. Pour l'équipe de direction, il peut conduire à engager ou à poursuivre la réflexion sur les moyens d'atteindre le minimum incompressi-

ble d'heures perdues, en envisageant les domaines où des actions relèvent de l'établissement et ceux où il subit des contraintes plus ou moins fortes. De plus, il révèle, sur un plan plus général, un certain nombre d'éléments connus de façon plus ou moins intuitive par tous les acteurs et observateurs du système scolaire français, et en propose une quantification.

Ainsi, les heures perdues pour cause de fermeture totale de l'établissement relèvent rarement de décisions de celui-ci et sont généralement incompressibles. Leur proportion peut être élevée et on constate que, pour les lycées qui peuvent fermer jusqu'à plusieurs semaines en fin d'année scolaire parce qu'ils sont centres d'examen, notamment pour le baccalauréat, elles constituent la première cause de pertes d'heures (une perte de 4.5 pour cent des heures annoncées aux élèves correspond à une fermeture moyenne de plus d'une semaine et demie dans ce type d'établissement). Si l'on y ajoute les 1.3 pour cent d'heures perdues parce que les enseignants sont mobilisés pour des préparations de sujets d'examen et concours, des jurys, des commissions diverses, etc., sans être remplacés, ce sont au total un peu plus de deux semaines, sur les 36 que dure théoriquement l'année scolaire, au cours desquelles l'enseignement ne peut être assuré pour des raisons tenant au fonctionnement interne du système (voir le tableau 10.5).

En ce qui concerne les heures perdues pour cause de formation continue des enseignants, on a, bien évidemment, tenu à les distinguer des précédentes et des suivantes dans la mesure où, si elles se traduisent, à court terme, par une perte d'enseignement par rapport à ce qui avait été prévu en début d'année scolaire, elles doivent, à plus long terme, être considérées comme un investissement pour le système éducatif.

Tableau/Table 10.5.

Proportion d'heures d'enseignement effectivement dispensées à un élève (F 2)

Références nationales pour 1992-93, issues d'une enquête
sur un échantillon national d'établissements

Student teaching hours actually received (F 2)

National reference data for 1992-93, from a survey on a nationwide sample of schools

Type d'établissement	Lycée	Lycées professionnels	Collèges
Heures d'enseignement prévues	100.0	100.0	100.0
Heures perdues pour :			
Fermeture totale de l'établissement	4.5	2.7	1.3
Absence non remplacées des enseignants			
• pour fonctionnement du système	1.3	1.8	1.1
• pour formation	1.2	2.9	0.5
• pour raisons individuelles	1.8	3.1	2.4
Absence de l'élève	3.2	4.6	2.6
Total heures perdues	12.0	15.1	7.9
Heures effectivement assurées	88.0	84.9	92.1

Les heures perdues par suite d'absences non remplacées des enseignants pour des raisons individuelles apparaissent relativement modestes par rapport à la place qu'elles peuvent tenir dans les discussions de certains conseils d'administrations de lycées et de collèges. Les enseignants sont, en tout état de cause, une des catégories professionnelles qui a, en France, un des taux d'absentéisme les plus bas. Plus fondamentalement, cet élément de l'indicateur doit conduire à s'interroger sur la capacité de remplacement des établissements. L'attribution à ceux-ci de moyens de remplacement est, en France, du ressort de l'autorité académique, mais la mise en œuvre d'un système de suppléance et la gestion des moyens attribués relèvent, quant à elles, de l'établissement.

Enfin, les heures perdues parce que les élèves sont absents peuvent être en nombre important. Si personne n'est surpris de constater que la proportion en est plus forte dans les lycées professionnels que dans les lycées d'enseignement général et technologique, et dans ceux-ci que dans les collèges, certains – par exemple, les parents – découvrent fréquemment, lorsque cet indicateur leur est présenté, que les élèves perdent, en moyenne, plus d'heures d'enseignement du fait de leurs propres absences que du fait du non-remplacement de leurs professeurs.

Cet indicateur se révèle donc particulièrement riche, tant pour le pilotage de l'établissement et l'éclairage des différents partenaires de la communauté éducative que pour les autorités académiques et nationales qui disposent ainsi de données homogènes qui devraient pouvoir, à terme, être prises en compte pour améliorer la gestion des ressources disponibles pour faciliter le remplacement des enseignants absents ou en formation. C'est sans doute, par ailleurs, un des indicateurs les plus délicats à présenter lors des formations données aux chefs d'établissement. Il est, en effet, nécessaire que ceux-ci aient bien à l'esprit ce que représentent les « heures dues à l'élève moyen du lycée », qui constituent la base de l'indicateur, sous peine de commettre des contresens.

Il s'agit de la durée annuelle de l'enseignement telle qu'elle a été prévue en début d'année dans l'établissement, c'est-à-dire le produit du nombre hebdomadaire moyen d'heures d'enseignement assurées à l'élève moyen du lycée et du nombre de semaines de cours, tel qu'il résulte du calendrier de l'année scolaire considérée. Cette donnée est, conceptuellement, différente de la durée annuelle théorique de l'enseignement dispensé aux élèves, proposée par l'un des indicateurs de *l'État de l'école en France* (ministère de l'Éducation nationale, 1993), et qui correspond à l'indicateur P 36 retenu par le Réseau C du projet INES. Cette durée annuelle théorique de l'enseignement est en effet établie à partir des horaires hebdomadaires prévus par les programmes officiels et l'on sait que, dans un établissement donné, la durée annuelle de l'enseignement effectivement prévue peut être (et est généralement) différente de cette durée théorique, en raison, par exemple, de l'organisation d'heures de renforcement ou de soutien dans certaines sections et disciplines, voire de l'impossibilité d'offrir l'horaire théorique d'une discipline, faute du poste et/ou de l'enseignant nécessaire. Il faut faire comprendre que ceci devra être pris en compte si l'on tente des comparaisons entre établissements ou dans le temps. Une même proportion d'heures « perdues » par les élèves ne se traduira pas par des résultats équivalents en termes de service rendu à ces élèves selon le caractère plus ou moins « copieux » de l'horaire hebdomadaire prévu à l'origine.

Par ailleurs, cet indicateur s'intéresse au temps d'enseignement dû à un élève, étant donné la structure pédagogique retenue dans l'établissement; cela le distingue de l'indicateur relatif aux moyens horaires (M 1) qui s'intéresse, lui, au temps d'enseignement dont dispose le collège pour l'ensemble des élèves. De plus, le non-remplacement d'un professeur absent pendant une heure a, pour l'indicateur, le même poids que l'absence d'un élève pendant une heure, quelle que soit la taille des groupes d'élèves auquel ce professeur enseigne.

L'importance des services liés à la vie scolaire

Cet indicateur (F 5) est présenté ici à titre d'exemple parmi ceux qui, construits selon le même procédé, ont pour ambition de décrire des aspects qualitatifs du fonctionnement des établissements.

Il mesure l'intensité des services liés à la vie scolaire par le biais des moyens dont dispose l'établissement, par l'usage qu'il en fait, ainsi que par les activités liées à la vie scolaire. L'importance de ces services reflète, dans une large mesure, la qualité du cadre dans lequel s'inscrit l'enseignement dispensé dans l'établissement.

L'indicateur distingue les ressources dont dispose l'établissement et les activités qui leur sont liées. Pour ce faire, il est établi à partir d'un ensemble de dix critères retenus à l'issue de discussions avec des chefs d'établissement :

- Pour apprécier l'importance des ressources dont dispose l'établissement :
 - nombre hebdomadaire d'heures de conseillers d'éducation¹/nombre d'élèves;
 - surface des locaux chauffés accessibles aux élèves à la pause méridienne/nombre d'élèves;
 - nombre annuel d'heures de présence du personnel médical et paramédical/nombre d'élèves;
 - nombre annuel d'heures de présence de l'assistante sociale/nombre d'élèves; et, pour les établissements ayant un internat :
 - surface totale de l'internat/nombre d'internes;
 - nombre d'internes de l'établissement bénéficiant d'un box ou d'une chambre individuelle/nombre d'internes de l'établissement.
- Pour apprécier l'importance des activités :
 - nombre d'élèves inscrits dans un ou plusieurs clubs²/nombre d'élèves;
 - nombre d'adhérents à l'association sportive de l'établissement/nombre d'élèves;
 - montant annuel des dépenses de l'établissement issues du fonds de la vie lycéenne ou du fonds d'animation³/nombre d'élèves;
 - nombre d'élèves ayant participé à au moins un projet d'action éducative⁴/nombre d'élèves.

Chacun de ces critères est assorti automatiquement d'une valeur de 1 à 5 selon la position qu'a sa valeur dans l'établissement par rapport à l'ensemble des valeurs d'un échantillon national d'établissements (de 1, si sa valeur dans l'établissement se situe parmi le cinquième de valeurs les plus faibles de l'échantillon national, à 5 si elle se situe parmi le cinquième de valeurs les plus fortes). Les valeurs ainsi obtenues sont ensuite

pondérées (grâce à une pondération retenue à l'issue des discussions avec des chefs d'établissement) pour établir une valeur globale de 1 à 5 qui donne une appréciation synthétique de l'importance des services offerts dans l'établissement et permet à celui-ci de se situer, par rapport aux autres, du point de vue de l'ensemble des critères retenus.

Le principe de construction de tels indicateurs multicritères, les critères retenus en définitive pour chacun d'entre eux et la pondération adoptée, font généralement l'objet de nombreuses discussions lors des formations des chefs d'établissement. Il faut leur montrer que les critères retenus ont moins d'intérêt considérés un par un, que pris ensemble et combinés pour rendre compte du phénomène qui nous intéresse. Il est utile, à ce sujet, d'indiquer qu'ils ont été choisis de façon à concilier deux soucis : celui de décrire de façon relativement complète la dimension du fonctionnement de l'établissement que veut apprécier l'indicateur et celui de ne retenir que des critères simples repérés par une information facile à constituer et à recueillir par l'établissement.

Ceci étant dit, on sait encore mal comment « fonctionneront » ces indicateurs qui n'ont été expérimentés pour l'instant que sur un nombre limité d'établissements.

4. Les premières réactions à la mise en place du dispositif et les perspectives

On ne peut, au moment où ce chapitre est écrit, que dresser un premier état de réactions et de réflexions issues, d'une part, des formations données aux chefs d'établissement avant qu'ils ne reçoivent le logiciel et la base standard d'indicateurs et, d'autre part, d'un premier bilan opéré auprès des quelque 600 chefs d'établissement qui les ont expérimentés en mai-juin 1994.

Premières réactions positives pour ce qui concerne l'outil d'auto-évaluation et de pilotage et quelques états d'âme par rapport à l'évaluation externe

Il est incontestable que la plupart des chefs d'établissement qui s'efforcent de se doter d'outils de pilotage et de compte rendu de la politique et du fonctionnement de leur établissement sont très intéressés. Ayant concrètement expérimenté les difficultés de la construction et de la pérennisation de quelques indicateurs, ils apprécient le service que représente la fourniture d'indicateurs tout calculés allant très au-delà de ce dont ils pouvaient disposer auparavant, et surtout d'indicateurs standardisés, comparables et assortis de références qui leur permettent de se situer par rapport aux autres établissements. De même, ils sont particulièrement sensibles au fait qu'une batterie standard d'indicateurs constituera une « mémoire » de l'établissement, immédiatement lisible par un chef d'établissement qui prend de nouvelles fonctions.

Il est, en tout cas, extrêmement frappant de voir comment les chefs d'établissements qui travaillent sur les cas pédagogiques du collège et du lycée qui leur sont proposés lors des formations, s'emparent des indicateurs pour établir un diagnostic très sévère de la politique conduite dans le collège et le lycée qui leur sont ainsi ouverts. Entre les mains de ces professionnels, l'outil se révèle extrêmement performant pour établir une première analyse de la situation d'un établissement et de la politique qui y est conduite et pour détecter des pistes d'action qui paraissent propres à améliorer cette politique.

C'est sans doute pourquoi certains chefs d'établissement manifestent quelque inquiétude devant les perspectives d'une évaluation externe de leurs établissements, et surtout d'eux-mêmes – la question de l'évaluation des chefs d'établissement étant, dans l'enseignement public secondaire français, une question rémanente depuis la fin de la décennie 80. Cependant, ceux qui sont inquiets d'une évaluation externe qui pourrait donner des « armes » aux responsables locaux de l'Éducation pour les juger, reconnaissent qu'un outil standardisé, aussi objectif que possible, permettant des comparaisons fondées et qui puisse être connu et maîtrisé par tous – aussi bien les chefs d'établissements que leur tutelle – devrait améliorer le dialogue avec cette tutelle et permettre à celle-ci de mieux jouer son rôle. On retrouve ici un des principes qui sont à la base de ce dispositif.

En tout état de cause, on ne peut qu'être satisfait de constater que, dès la première mise en place du dispositif, celui-ci a, pour tout ou partie, déjà fait l'objet de premières utilisations dans quelques établissements en permettant, par exemple, de nourrir le rapport du chef d'établissement au Conseil d'administration de fin d'année (ce qui a, dans certains cas, conduit le conseil d'administration à revoir quelques idées reçues sur l'établissement ou la population qu'il accueille). De plus, certains responsables locaux de l'Éducation nationale aménagent, voire modifient, le contenu des comptes rendus qu'ils demandent aux établissements et en renvoient le calendrier afin qu'ils soient cohérents avec le dispositif « Indicateurs pour le pilotage des établissements secondaires ».

Enfin, en ce qui concerne la diffusion publique d'indicateurs relatifs aux établissements, un pas important vient d'être franchi avec la publication par la Direction de l'évaluation et de la prospective du ministère de l'Éducation nationale de trois indicateurs de performance, lycée par lycée, publication qui concerne 2 143 lycées publics et privés sous contrat préparant au baccalauréat général et technologique et 1 487 lycées professionnels publics et privés sous contrat préparant au baccalauréat professionnel, c'est-à-dire la totalité d'entre eux, à l'exception de quelques-uns qui ne présentent à l'examen qu'une poignée de candidats. Il s'agit là d'une pratique tout à fait nouvelle pour l'Éducation nationale française; si elle suscite des interrogations, il semble que l'on apprécie plutôt dans les établissements le fait que des données objectives sur ce sujet délicat soient publiées sous l'égide du ministère, ce qui devrait faire taire les rumeurs qui circulent sur les établissements et les publications parfois mal fondées de certains organes de presse.

Des perspectives d'évolution de plusieurs ordres

En ce qui concerne l'évolution de l'outil lui-même, deux voies de progrès devraient être rapidement explorées.

Il s'agit, tout d'abord d'une amélioration des références fournies avec les indicateurs standards. Les chefs d'établissement sont très demandeurs de références correspondant à des établissements comparables au leur, et pas seulement de références nationales ou de références établies sur des aires géographiques ou administratives. L'élaboration de typologies d'établissements sur des bases plus riches que celles établies jusqu'ici sera un thème majeur des programmes de travail des années à venir. C'est certainement une nécessité pour permettre aux établissements de mieux se situer, mais également pour que les appréciations externes soient mieux étayées.

Il s'agit ensuite d'un enrichissement de la batterie standard d'indicateurs. La création de nouveaux indicateurs a déjà été évoquée : intensité des relations avec les parents, importance de la formation continue du personnel, importance des activités de formation des adultes et de promotion sociale dans les établissements, par exemple. Cependant, s'il est incontestable qu'il convient de compléter cette batterie – ce que se proposent d'ores et déjà de faire certaines académies, en particulier en fournissant aux établissements des indicateurs de résultats issus de cohortes réelles de leurs élèves – il sera sans doute nécessaire, à tous les niveaux (central, académique et local) de résister à la tentation de transformer une batterie synthétique d'indicateurs en une base de données statistiques sur l'établissement.

En ce qui concerne l'utilisation de l'outil, on perçoit encore mal son intérêt lorsque les établissements disposeront d'une évolution dans le temps de leurs indicateurs. Si, pour l'instant, l'analyse synchronique domine dans les formations et les premières utilisations, l'analyse diachronique devrait, au fil des ans, prendre une place au moins équivalente.

Dans le même ordre d'idées, il reste à explorer et à faire apprécier lors des formations, tant initiales que continues des chefs d'établissement, des inspecteurs et des responsables locaux, en quoi la mise à disposition de cet outil nouveau doit modifier et faire évoluer les conditions d'exercice de leurs métiers respectifs.

Enfin, il est bien évident que la possibilité de disposer, au plan national, de données nouvelles sur le fonctionnement des collèges et des lycées devrait permettre d'améliorer la connaissance que l'on a des établissements secondaires français et des caractéristiques de leur fonctionnement et constituer une base d'études originales qui permettra de faire évoluer le dispositif d'indicateurs lui-même.

Remerciements

Cet article rend compte d'un travail collectif qui s'est étendu sur plusieurs années ; il doit beaucoup à tous ceux qui y ont participé, en particulier Philomène Abi Saab, Thérèse Elhachem, Sabine Meunier, Marie-Claude Rondeau, chargées d'études au Département de l'évaluation des systèmes et des unités d'éducation et Claude Sauvageot, chargé de mission à la Direction de l'évaluation et de la prospective, ainsi, bien sûr, qu'à Denis Meuret, qui a dirigé le Département de l'évaluation des systèmes et des unités d'éducation jusqu'à la rentrée scolaire 1993.

Notes

1. Les conseillers d'éducation assurent dans les établissements du second degré les tâches d'organisation de la vie scolaire, d'éducation et de surveillance.
2. Les clubs proposent aux élèves, en dehors des heures d'enseignement, des activités très diverses. Ils sont organisés sur la base du volontariat par des élèves, des enseignants ou d'autres adultes.
3. Ces fonds sont consacrés à des dépenses décidées par les délégués des élèves.
4. Ces projets, de nature très diverses, consistent en l'organisation d'activités en lien direct ou non avec l'enseignement et la vie scolaire : par exemple, ateliers de lecture, organisation de spectacles, etc.

Références

- FIXARI, D. et KLETZ, F. (1993), *Où en est l'auto-évaluation des établissements scolaires ?*, Rapport d'étude pour la Direction de l'Évaluation et de la Prospective, ministère de l'Éducation nationale, Paris.
- Inspection générale de l'Éducation nationale, ministère de l'Éducation nationale (1991), *Rapport de l'Inspection générale de l'Éducation nationale*, La Documentation française, Paris.
- Inspection générale de l'Éducation nationale, Inspection générale de l'Administration de l'Éducation nationale, ministère de l'Éducation nationale (1993, décembre), *Rapport général sur les évaluations d'établissement 1989-1992*, ministère de l'Éducation nationale, Paris.
- LIENSOL, B. et MEURET, D. (1987), « Les performances des lycées publics et privés pour la préparation au baccalauréat », *Éducation et Formations*, n° 11, ministère de l'Éducation nationale, Paris.
- Loi d'orientation sur l'éducation, Loi n° 89-486 du 10 juillet 1989, *Journal officiel de la République française* du 14 juillet 1989, Paris.
- Ministère de l'Éducation nationale (1993), *L'État de l'École*, ministère de l'Éducation nationale, Direction de l'évaluation et de la prospective, n° 3, Paris, octobre.
- Ministère de l'Éducation nationale (1994a), *Géographie de l'École*, ministère de l'Éducation nationale, Direction de l'évaluation et de la prospective, n° 2, Paris, janvier.
- Ministère de l'Éducation nationale (1994b), *Indicateurs pour le pilotage des établissements du second degré (IPES)*, Guides de l'utilisateur, ministère de l'Éducation nationale, Direction de l'évaluation et de la prospective, mars.
- Ministère de l'Éducation nationale (1994c), *Trois indicateurs de performance des lycées, Baccalauréat 1993, Résultats lycée par lycée*, Les dossiers d'Éducation et Formations, n° 41, tomes I et II, ministère de l'Éducation nationale, Paris, juin.
- RONDEAU, M.-C. (1990), « Des indicateurs statistiques à l'usage des établissements du second degré », *Éducation et Formations*, n° 22, ministère de l'Éducation nationale, Paris.
- THÉLOT, C. (1991), « Mesurer l'efficacité des lycées », *Le Monde de l'Éducation*, n° 179.
- THÉLOT, C. (1993), *L'évaluation du système éducatif*, Nathan, Paris.
- VERT, E. (1990), « Les performances des lycées au baccalauréat », *Éducation et Formations*, n° 22, ministère de l'Éducation nationale, Paris.

Chapter/Chapitre 11

Interpretation and Use of Education Indicators
Interprétation et exploitation des indicateurs de l'enseignement

by

Erik Wallin

National Agency for Education, Stockholm, Sweden

Most OECD countries have during recent decades experienced a change in the relations between different levels of the education system. This reflects a change in the role of the state, and a need for changed methods and models of steering and managing the education system. Together with this development goes a change in the knowledge expected from educational research and from evaluative activities. As regards educational research, the demand is less for instrumental answers to educational issues than for perspectives and conceptual clarifications. There is a need for knowledge that leads to increased understanding of educational phenomena as a basis for action. The same is true of knowledge generated by evaluative activities. Outcome data are of interest, but so are process data, which lead to an understanding of how results come about. This is of vital interest in a new situation of guidance by goals and results. Both the school and the decision-makers at the municipal and state levels should know about results and how they have come about in order to be able to take necessary steps for change and development. To a considerable extent, the purpose of Network C indicators is to offer precisely that kind of data, which can make data from the other networks understandable. The purpose of this chapter is to discuss some of the implications for Network C.

*

* *

Note de synthèse

Au cours des dernières décennies, il s'est produit dans la plupart des pays de l'OCDE un changement des rapports entre les divers niveaux du système d'enseignement. Ce changement reflète pour l'essentiel un changement du rôle de l'État et induit la nécessité de modifier les moyens et les modèles de direction et de gestion du système d'enseignement.

Parallèlement à cette évolution, on voit changer les attentes quant à l'information que l'on peut tirer de la recherche pédagogique et des activités d'évaluation. Pour ce qui est de la recherche pédagogique, on attend des réponses utilitaires aux problèmes qui se posent dans l'enseignement mais surtout de nouveaux points de vue et des clarifications conceptuelles. On espère disposer de connaissances qui offrent des éclairages nouveaux, autrement dit, des connaissances qui permettent de mieux comprendre les phénomènes de l'enseignement afin d'agir en conséquence.

Il en est de même des connaissances issues des activités d'évaluation. Les données relatives aux résultats sont certes intéressantes, mais celles relatives aux processus le sont aussi lorsqu'elles débouchent sur la compréhension de la façon dont ces résultats ont été obtenus. Cette information présente un intérêt vital dans la nouvelle situation créée par la direction selon les objectifs et résultats. L'école et les décideurs qui agissent aux niveaux municipal, régional ou national doivent tous être informés des résultats et de la façon dont ils sont advenus afin de pouvoir prendre les indispensables mesures en faveur du changement et du développement.

Dans une très grande mesure, les indicateurs du Réseau C ont précisément pour but de fournir les types de données qui rendent compréhensibles les données provenant des autres réseaux.

Ce chapitre a pour objet d'étudier les incidences de cette situation pour le réseau.

*

* *

1. Introduction

In most of the OECD countries there is a trend towards restructuring the education system, and particularly its control and management. In some countries, such as Sweden, there is much talk about decentralisation, and in others, such as the United States, of tendencies towards centralisation. Centralisation and decentralisation seem to be rather meaningless concepts as they are relative: although Sweden is decentralised it is still more centralised than the United States would be, even if that country were to acquire a more centralised system. The common denominator is *restructuring*, a *change of relationships* between different levels of the education system, or *changes in the role of the state*. In another sense, this is leading to the need for changed methods and models of steering and managing the education system.

Together with this development goes a change in the expectations of the knowledge to be generated by educational research. Instead of instrumental answers to educational issues, the demand is more for perspectives and conceptual elaborations. The desired answers are more oriented towards further clarification of the research problem than of a contextually barren solution.

A main reason for this state of affairs is a change in the conception of teachers and teaching, but also the notion that with changed local and professional responsibility, policy-making and local curriculum work are acquiring a different focus. There is a need for knowledge with what has been termed an "enlightenment function": that is, knowledge from both quantitative and qualitative research that leads to increased understanding as a basis for action.

The same is true of evaluation and knowledge generated by evaluative activities. There is a need for data that not only provide educational results but also lead to an understanding of how results come about through educational processes. This is of vital interest in the new situation of decentralisation and orientation towards goals. Both the self-evaluating school and the decision-maker at a municipal and state level should know about results and how they have come about in order to be able to take necessary steps for change and development.

To some extent, the purpose of Network C indicators is to offer the kind of data that can make data from other networks better understandable. The purpose of this essay is to discuss some of the implications for the network.

2. Network C – A Context for Interpretation

According to its remit and title of "the functioning of schools", Network C has chosen to work with indicators of two different kinds. One kind describes inputs or prerequisites for the functioning of schools, such as teacher education, teachers' salaries, pupil : teacher ratio and modes and locus of decision-making. Underlying this, in an explanatory perspective, come public expenditures on education. This means that there are prerequisites that are chained to each other but not, which seems important, in a linear fashion. This is one way of saying that the whole endeavour is an issue of understanding.

Another kind of indicator consists of data concerning the educational processes taking place at the school level in terms of work organisation, models of communication and teaching methods. These indicators are, among others, differentiation and individualisation, evaluation practices, and school leadership.

As stated in the above, different indicators go together in a complicated relationship. Naming indicators input or process indicators is therefore just a matter of labelling and not of definition in a more formal way. If a dichotomy (prerequisites and processes) is accepted, there are also indicators that fall in between. One example of these is the indicator, of great importance, concerned with the intended curriculum. This indicator describes the outline of subjects for different International Standard Classification of Education (ISCED) levels and the time given to them.

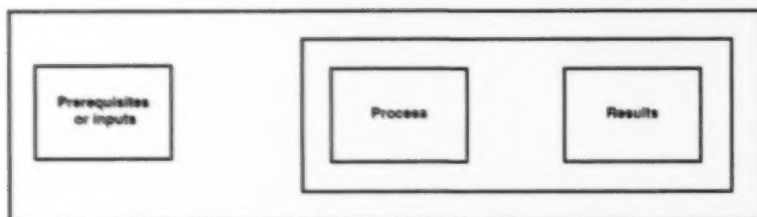
Both input and process indicators have a value and meaning in themselves, particularly the latter. They illustrate the ambitions of Network C, which focus on understanding and interpretation. It is common to the Network C indicators that they are intended, primarily, to make possible a description of the contextual background against which product data from, for instance, Network A (*i.e.* the network within the INES project concerned with output and outcome indicators) can be understood and interpreted in a comparative perspective.

The intention of this chapter is to underline the function of Network C indicators in the interpretation of indicators from other networks and hence also of the whole INES project. It should, however, be remarked that some process indicators, such as models for differentiation of pupils and co-operation among staff in schools, should be considered as results in themselves. Suffice it here to say that this is because observations of outcomes of education in many of its most important aspects (such as social development and learning to solve problems) are likely to be made in terms of processes which, it is assumed, will lead to certain effects.

Figure 11.1 illustrates the relationships discussed: prerequisites or inputs and processes are related to each other, and these two together are related to results. The outer surrounding line in Figure 11.1 indicates the necessity of a holistic approach to understanding and interpretation. This presentation symbolises the possibility of going beyond linear models of interpretation of the relations between inputs, processes and outputs.

One consequence of this way of looking at the ambitions of Network C will be the close relationship with Network A, which has initiated work on the issue of goals. These, to make the picture complete, should be included in Figure 11.1. A main reason for this is that results cannot be understood and comparisons cannot be made without reference to the intentions and emphases coded in the goals. Another reason is that considering outcomes of an education system and making any kind of comparison (*e.g.* country comparisons) without considering the objectives, will run the risk of contributing to a

Figure 11.1. A revised holistic model relating inputs to processes and results
Modèle global révisé concernant les éléments intervenant dans les processus et les résultats



Source: OECD.

static system. The outcomes will be without meaning in relation to intentions and are thus open to analysis and discussion in relation to any criteria, or rather to no criteria beyond the analyst's own ideosyncratic ideas. The definition of criteria runs into the danger that it may be governed by existing schools of thought rather than by what is good or bad, as defined by the agreed goals.

So, in considering goals, Network A seems to be approaching the field of Network C, which calls for co-operation between the two networks. As a matter of fact, this is also a recommended course of action for all networks of INES, especially in these final years of the developmental period. A total view of the different INES networks and their related outcomes could form the base for a more analytic and discursive publication from the project as a complement to *Education at a Glance* (OECD, 1995).

3. INES – Its Context for Use of Data

So far, Network C and its work in relation to the other INES networks has been considered as an issue in itself. How can this discussion be related to the introductory remarks on restructuring and the new knowledge needs of the education system?

Restructuring of the Education System

The restructuring of the education system taking place in different countries is intended to provide greater freedom and flexibility to municipalities, schools and teachers in relation to organisation of schools and to selection, sequencing and structuring of knowledge and learning activities. Needless to say, this whole process takes place within frameworks set by the state. A concept of curriculum that only focuses on state guidance through statements of intentions, may be appropriate in the context of guidance by regulation. This is not the case if goals are the guiding force. At a time when the responsibility for organisation and educational activities resides at the level of schools and teachers, what ultimately happens to the state's intentions is determined by local decision-makers and teachers and students in the classrooms. In addition, the potential for differences between what the state intends and what is actually offered by the municipalities to pupils and parents is likely to increase. For these reasons, it is important that there should exist a vital relationship between state and municipalities and that any conception of "curriculum" should include the dynamic relationship between state intentions and the interpretation and realisation of these in classrooms. A main task for local politicians is to provide the necessary and appropriate prerequisites for the schools' work and for the teachers to master the local curriculum work.

All knowledge production within the educational field should observe the demands implicit in the situation so described. Use of knowledge for decision-making is being given new emphasis for politicians and civil servants as well as for teachers and school principals. The latter are all acquiring new or expanded roles as providers of opportunities for education and curriculum work. The decision-making is up to them. They are not expected to accept ready-made solutions but to take advantage of the expanded possibilities given by the state to use knowledge to make good decisions.

The INES project should also be seen in this perspective. Its main function is to offer data to be interpreted and used for different purposes at different levels of the state and the education system. As the project addresses varied aspects of the education system, as well as contextual questions, its output can be used in different ways. The reception of the volumes of *Education at a Glance* which are already published has proved this.

Decision-Makers and Teachers as Decision-Makers

David Hamilton (in a study for the Swedish Agency for Education) says of teacher education and training that it should not equip the prospective teacher with a knowledge of absolute laws on learning mechanisms or with ready-made methodological solutions (Hamilton, 1992). Instead, it should prepare the teacher-to-be with "indeterminate knowledge" to be used in "indeterminate settings". Knowledge production related to this has as its prime task to ask the necessary questions and in this way to provide the future teacher with a basis also for asking the relevant questions. An implicit understanding of teaching implicitly based on a "research and development" (R&D) model has been replaced with an understanding of teaching as action based on knowledge which permits reflection and individual decisions. Thus, there is a conception of teaching and the teacher that emphasises the decision-making function.

Knowledge production for the new guidance environment – and thus for knowledge production of the INES kind – is, of course, not only directed to teachers and school principals but also, and more importantly, to policy-makers and educational planners. But the underlying principles of teaching are also applicable to policy-making and educational planning. For a long time, decision-making as policy-making and planning has been thought of as a linear, rational process in which technical solutions and directives are balanced against each other and against pre-defined criteria. Research in recent years has emphasised planning and policy-making as a process dominated by uncertainty, in which decisions emerge gradually and in which knowledge serves to provoke questions and to create contexts and overviews. That type of rationality is not one of ends and means, but of what is appropriate and feasible in a given context. Two of the leading researchers in this field (Weiss and Boucouvalas, 1977) have spoken of the "enlightenment function" of research in the social sciences. Planning and policy-making, just like teaching, have the character of action. Knowledge can, quite simply, make policy-makers and planners wiser.

Reports on INES indicator data have the same function, namely to provide knowledge to be used for decision-making. Thus, results should not be considered to have meanings of their own but instead to obtain their meaning in the context of the totality of the indicators. This is to stress the enlightenment function, but it is also to reduce the danger that isolated data may be interpreted only in relation to the criteria set up by the reader, or, in other words, to decrease the number of degrees of freedom of interpretation.

The Encounter between Knowledge and User

All experience, according to common observation as well as to research, conveys a relatively uniform picture as far as knowledge utilisation and research information are

concerned: these are difficult areas, further measures are needed, and the "more-of-the-same" approach is in need of revision. The implications of a R&D model, which is based on the diffusion of knowledge, must be replaced by a mentality centred on the user and communication. The R&D model in its purest form implies the production of knowledge in one context for use in another. The transmission problem is often regarded as a quantitative problem of dissemination, the solution being to popularise. This is to disregard the various cultural and communicative preconditions to which, quite clearly, the different environments are subject. The research community, and other knowledge-producing bodies, are characterised by their preconditions and purposes, as are the policy-making and teaching community and the sectors that are expected to "apply" the new knowledge. For certain types of decision, it is clearly the case that research findings can be translated into action relatively directly and easily. In the education sector, the problems are generally more complicated, and the disappointing impact of a mentality which relies on the diffusion of results is not surprising.

This being so, the provision of information is a matter of creating arenas for communication on the subject of knowledge production and its results. These arenas provide opportunities for discussion in which experience-based knowledge can be confronted with knowledge derived from research or evaluation. This is to provide opportunities for encounters between different interests, needs, cultures and conceptions of the field being studied in one form or another. In other words, it is an encounter between different perspectives and different frames of reference, out of which comes a new understanding on all sides. Or, it can be expressed as a matter of creating new meaning. In the educational field, it is an encounter between the practical experience of, say, teachers and policy-makers, and the scientifically derived knowledge of the researcher. This should be seen as an encounter between different parties on equal terms and not in a hierarchical fashion.

4. Back to INES Data

As has already been said, the intention of this chapter is to underline the importance of an interpretative use of INES data. The discussion so far has, however, not been confined to the kind of data offered by the INES project but has more general implications in terms of the uses of data and knowledge produced by different knowledge-generating actors, primarily in research and evaluation. Figure 11.1, which summarises an analytic standpoint, contains some general relations that have to be theoretically and conceptually elaborated for any particular case in order to be analytically generative.

When used as a general directive for analysing INES data, however, the model shown in Figure 11.1 has a number of possibilities:

- It can be considered by the individual reader when reading *Education at a Glance*.
- It can be considered as giving directives for structuring *Education at a Glance* and binding its different parts together.
- It can be used to formulate comments on *Education at a Glance* and on an interpretative reading of the content.

- As a main purpose of the INES project is to standardise and present education statistics across a range of countries, it is easy to argue that *Education at a Glance* should just present statistics and no interpretative comments. If this is the case, consideration could be given to producing a supplement to *Education at a Glance* presenting a more interpretative picture of results. (This could also be part of the further development of the system of INES indicators, as the analytical work would pinpoint difficulties and needs for improvements.)
- It could encourage countries to produce separate (or bilateral, trilateral, etc.) reports focusing on comparisons and interpretations in order to achieve a sound basis for policy-making.

The summary message of this chapter is that INES has to encourage good use of data and avoid its effects becoming static or reactionary. Use of knowledge should be problem-oriented, and it is important to find a structure for organising indicators. As a matter of fact, the chapter also asks what an indicator is. It advocates the use of INES data in context and not as a number of isolated facts that have little meaning in themselves. Network C has a special responsibility for achieving this.

References

- HAMILTON, D. (1992), *Steering the Future: Rationality and Efficiency in Policy-Related Educational Research and Development*, Mimeo, National Agency for Education, Stockholm.
- WEISS, C. and BOUCOUVALAS, M.J. (1977), *Social Science and Decision Making*, Columbia University Press, New York.

ALSO AVAILABLE

Education at a Glance - OECD Indicators

FF 220 FFE 285 £35 US\$ 54 DM 83

OECD Education Statistics, 1985-1992/Statistiques de l'enseignement de l'OCDE, 1985-1992 (bilingual)

FF 160 FFE 210 £25 US\$ 40 DM 60

Measuring What Students Learn/Mesurer les résultats scolaires (bilingual)

FF 110 FFE 140 £17 US\$ 27 DM 40

Education and Employment/Formation et emploi (bilingual)

FF 90 FFE 115 £14 US\$ 22 DM 34

Public Expectations of the Final Stage of Compulsory Education/Le dernier cycle de l'enseignement obligatoire : quelle attente ? (bilingual)

FF 120 FFE 155 £20 US\$ 29 DM 47

Decision-Making Processes in the Education Systems of 14 OECD Countries (forthcoming)

Prices charged at the OECD Bookshop.

The OECD CATALOGUE OF PUBLICATIONS and supplements will be sent free of charge on request addressed either to OECD Publications Service, or to the OECD Distributor in your country.

ÉGALEMENT DISPONIBLES

Regards sur l'éducation - Les indicateurs de l'OCDE

FF 220 FFE 285 £35 US\$ 54 DM 83

OECD Education Statistics, 1985-1992/Statistiques de l'enseignement de l'OCDE, 1985-1992 (bilingue)

FF 160 FFE 210 £25 US\$ 40 DM 60

Measuring What Students Learn/Mesurer les résultats scolaires (bilingue)

FF 110 FFE 140 £17 US\$ 27 DM 40

Education and Employment/Formation et emploi (bilingue)

FF 90 FFE 115 £14 US\$ 22 DM 34

Public Expectations of the Final Stage of Compulsory Education/Le dernier cycle de l'enseignement obligatoire : quelle attente ? (bilingue)

FF 120 FFE 155 £20 US\$ 29 DM 47

Les processus de décision dans 14 systèmes éducatifs de l'OCDE (à paraître prochainement)

*Prix de vente au public dans la librairie du siège de l'OCDE.
LE CATALOGUE DES PUBLICATIONS de l'OCDE et ses suppléments seront envoyés
gratuitement sur demande adressée soit à l'OCDE, Service des Publications,
soit au distributeur des publications de l'OCDE de votre pays.*

MAIN SALES OUTLETS OF OECD PUBLICATIONS PRINCIPAUX POINTS DE VENTE DES PUBLICATIONS DE L'OCDE

ARGENTINA - ARGENTINE

Carlos Hirsch S.R.L.
Galeria Guernica, Florida 165, 4° Piso
1333 Buenos Aires Tel: (1) 331 1787 y 331 2391
Telefax: (1) 331 1787

AUSTRALIA - AUSTRALIE

D.A. Information Services
648 Whitmore Road, P.O. Box 163
Mitcham, Victoria 3132 Tel: (05) 873 4411
Telefax: (05) 873 5679

AUSTRIA - AUTRICHE

Gerold & Co
Graben 31
Wien 1 Tel: (0222) 533 50 14

BELGIUM - BELGIQUE

Jean De Lancy
Avenue du Roi 202
B-1060 Bruxelles Tel: (02) 538 51 69/538 08 41
Telefax: (02) 538 08 41

CANADA

Renouf Publishing Company Ltd
1294 Algoma Road
Ottawa, ON K1B 3W8 Tel: (613) 741 4533
Telefax: (613) 741 5439

Stones

61 Sparks Street
Ottawa, ON K1P 5R1 Tel: (613) 238 8985
211 Yonge Street
Toronto, ON M5B 1M4 Tel: (416) 363 3171
Telefax: (416) 363 59 63

Les Editions La Liberté Inc.

3020 Chemin Sainte-Foy
Sainte-Foy, PQ G1X 3V6 Tel: (418) 658 3763
Telefax: (418) 658 3763

Federal Publications Inc.

165 University Avenue, Suite 701
Toronto, ON M5H 1B8 Tel: (416) 860 1611
Telefax: (416) 860 1608

Les Publications Fédérales

1185 Université
Montréal, QC H3B 3A7 Tel: (514) 954 1633
Telefax: (514) 954 1633

CHINA - CHINE

China National Publications Import
Export Corporation (CNPIEC)
16 Gongti E. Road, Chaoyang District
P.O. Box 88 or 90
Beijing 100704 PR Tel: (01) 506 6608
Telefax: (01) 506 3101

CZECH REPUBLIC - RÉPUBLIQUE TCHÈQUE

Arna Pegas Press Ltd
Narodni Trida 25
POB 825
111 21 Praha 1 Tel: 26 65 68
Telefax: 26 20 81

DENMARK - DANEMARK

Munksgaard Book and Subscription Service
35, Nørre Søgade, P.O. Box 2148
DK-1016 København K Tel: (33) 12 85 70
Telefax: (33) 12 93 87

EGYPT - ÉGYPTE

Middle East Observer
41 Sherif Street
Cairo Tel: 392 6919
Telefax: 360 6804

FINLAND - FINLANDE

Akametsen Kirjakauppa
Keskuskatu 1, P.O. Box 128
00100 Helsinki
Subscription Services/Agrégat d'abonnements
P.O. Box 23
00071 Helsinki Tel: (358 0) 12141
Telefax: (358 0) 121 4450

FRANCE

OECD/OCDE
Mail Orders/Commandes par correspondance
2, rue André-Pascal
75775 Paris Cedex 16 Tel: (33-1) 45 24 82 00
Telefax: (33-1) 49 10 42 76
Telex: 680048 OCDE

Orders via Minsat, France only/
Commandes par Minsat, France exclusivement
36 15 OCDE

OECD Bookshop/Librairie de l'OCDE
33, rue Octave-Fouillet
75016 Paris Tel: (33-1) 45 24 81 67
(33-1) 45 24 81 81

Documentation Française

29, quai Voltaire
75007 Paris Tel: 40 15 70 00

Gibert Jeune (Droit-Economie)

6, place Saint-Michel
75006 Paris Tel: 43 25 91 19

Librairie du Commerce International

10, avenue d'Alma
75016 Paris Tel: 40 73 34 60

Librairie Dunod

Université Paris-Dauphine
Place du Maréchal de Lattre de Tassigny
75016 Paris Tel: (1) 44 05 40 13

Librairie Lavoisier

11, rue Lavoisier
75008 Paris Tel: 42 65 39 95

Librairie L.G.D.J. - Monchrestien

20, rue Soufflot
75005 Paris Tel: 46 33 89 85

Librairie des Sciences Politiques

36, rue Saint-Guillaume
75007 Paris Tel: 45 48 36 02

P.U.F.

48, boulevard Saint-Michel
75005 Paris Tel: 43 25 83 40

Librairie de l'Université

12a, rue Nassef
13100 Aix-en-Provence Tel: (16) 42 26 18 08

Documentation Française

165, rue Garibaldi
69003 Lyon Tel: (16) 78 63 32 23

Librairie Decize

29, place Bellecour
69002 Lyon Tel: (16) 72 40 54 54

GERMANY - ALLEMAGNE

OECD Publications and Information Centre
August-Bebel-Allee 6
D-53175 Bonn Tel: (0228) 959 120
Telefax: (0228) 959 12 17

GREECE - GRÈCE

Librairie Kauffmann
Marokordanos 9
106 78 Athens Tel: (01) 32 55 321
Telefax: (01) 36 33 967

HONG-KONG

Sunderland Book Co. Ltd
13-15 Lock Road
Kowloon, Hong Kong Tel: 2376 2062
Telefax: 2376 0685

HUNGARY - HONGRIE

Euro Info Service
Margitsziget, Europa Ház
1138 Budapest Tel: (1) 111 62 16
Telefax: (1) 111 60 61

ICELAND - ISLANDE

M&M Mink Mennings
Laugavegi 18, Pósthólfur 192
121 Reykjavik Tel: 162 35 23

INDIA - INDE

Oxford Book and Stationery Co
Sindia House
New Delhi 110001 Tel: (11) 331 5896/5108
Telefax: (11) 332 5993

17 Park Street
Calcutta 700016 Tel: 240832

INDONESIA - INDONÉSIE

Pili-Lipi
P.O. Box 4298
Jakarta 12042 Tel: (21) 573 34 67
Telefax: (21) 573 34 67

IRELAND - IRLANDE

Government Supplies Agency
Publications Section
4/5 Harcourt Road
Dublin 2 Tel: 661 31 11
Telefax: 661 06 45

ISRAEL

Pradista
5 Shattar Street
P.O. Box 34030
Jerusalem 91430 Tel: (2) 52 84 90/1/2
Telefax: (2) 52 84 93

ROY

P.O. Box 13056
Tel Aviv 61130 Tel: (3) 49 61 08
Telefax: (3) 544 60 39

ITALY - ITALIE

Libreria Commissionaria Sansoni
Via Duca di Calabria 1/1
50125 Firenze Tel: (055) 64 54 15
Telefax: (055) 64 12 57

Via Bolchini 29
20155 Milano Tel: (02) 36 50 83

Editrice e Libreria Herder
Piazza Montecitorio 120
00186 Roma Tel: 679 46 28
Telefax: 678 47 51

Libreria Hoepli
Via Hoepli 5
20121 Milano Tel: (02) 86 54 46
Telefax: (02) 805 28 86

Libreria Scientifica
Dott. Lucio de Busto "Aesou"
Via Coronelli, 6
20146 Milano Tel: (02) 48 95 45 52
Telefax: (02) 48 95 45 48

Libreria Hoepli
Via Hoepli 5
20121 Milano Tel: (02) 86 54 46
Telefax: (02) 805 28 86

Libreria Scientifica
Dott. Lucio de Busto "Aesou"
Via Coronelli, 6
20146 Milano Tel: (02) 48 95 45 52
Telefax: (02) 48 95 45 48

Libreria Hoepli
Via Hoepli 5
20121 Milano Tel: (02) 86 54 46
Telefax: (02) 805 28 86

Libreria Scientifica
Dott. Lucio de Busto "Aesou"
Via Coronelli, 6
20146 Milano Tel: (02) 48 95 45 52
Telefax: (02) 48 95 45 48

JAPAN - JAPON

OECD Publications and Information Centre
Landis Akasaka Building
2-3-4 Akasaka, Minato-ku
Tokyo 107 Tel: (81 3) 3586 2016
Telefax: (81 3) 3584 7929

KOREA - CORÉE

Kyobo Book Centre Co. Ltd
P.O. Box 1658, Kwang Hwa Moon
Seoul Tel: 730 78 91
Telefax: 735 00 30

MALAYSIA - MALAISIE

University of Malaya Bookshop
University of Malaya
P.O. Box 1127, Jalan Pantai Baru
59100 Kuala Lumpur
Malaysia Tel: 756 5000/756 5425
Telefax: 756 5246

MEXICO - MEXIQUE

Revistas y Periodicos Internacionales S.A. de C.V.
Floresca 57 - 1004
Mexico, D.F. 06600 Tel: 207 81 00
Telefax: 208 39 79

NOT COPY AVAILABLE

NETHERLANDS - PAYS-BAS

SDU Uitgeverij Plantijnstraat
Ekinne Fondsen
Postbus 20014
2500 EA 's-Gravenhage
Voor bestellingen: Tel: (070) 37 89 880
Telefax: (070) 34 75 778

**NEW ZEALAND
NOUVELLE-ZÉLANDE**

Legislation Services
P.O. Box 12418
Thorndon, Wellington
Tel: (04) 496 5632
Telefax: (04) 496 5698

NORWAY - NORVÈGE

Narvesen Info Center - NIC
Bertrand Narvesen vei 2
P.O. Box 6125 Etterstad
0602 Oslo 6
Tel: (022) 57 33 00
Telefax: (022) 68 19 01

PAKISTAN

Mirza Book Agency
65 Shahrah Quaid-E-Azam
Lahore 54000
Tel: (42) 353 601
Telefax: (42) 231 730

PHILIPPINE - PHILIPPINES

International Book Center
5th Floor, Filipinas Life Bldg
Ayala Avenue
Metro Manila
Tel: 81 96 76
Telefax: 23312 RHP PH

PORTUGAL

Livraria Portugal
Rua do Carmo 70-74
Apart. 2681
1200 Lisboa
Tel: (01) 347 49 82/5
Telefax: (01) 347 02 64

SINGAPORE - SINGAPOUR

Gower Asia Pacific Pte Ltd.
Golden Wheel Building
41, Kallang Pudding Road, No. 04-03
Singapore 1334
Tel: 741 5166
Telefax: 742 9356

SPAIN - ESPAGNE

Mundi-Prensa Libros S.A.
Castelló 37, Apartado 1223
Madrid 28001
Tel: (91) 431 33 99
Telefax: (91) 575 39 98

Libreria Internacional AEDOS

Consejo de Censo 391
08009 - Barcelona
Tel: (93) 488 30 09
Telefax: (93) 487 76 59

Libreria de la Generalitat

Palau Moja
Rambla dels Estudis, 118
08002 - Barcelona
(Subscriptions) Tel: (93) 318 80 12
(Publications) Tel: (93) 302 67 23
Telefax: (93) 412 18 54

SRI LANKA

Centre for Policy Research
c/o Colombo Agencies Ltd
No. 300-304, Galle Road
Colombo 3
Tel: (1) 574240, 575551-2
Telefax: (1) 575394, 510711

SWEDEN - SUÈDE

Frizes Information Center
Box 16356
Regeringsgatan 12
106 47 Stockholm
Tel: (08) 690 90 90
Telefax: (08) 20 50 21

Subscription Agency/Agence d'abonnements
Wentweges-Williams Info AB
P.O. Box 1305
171 25 Solna
Tel: (08) 705 97 50
Telefax: (08) 27 00 71

SWITZERLAND - SUISSE

Madrac S.A. (Books and Periodicals - Livres
et périodiques)
Chemin des Palentes 4
Case postale 266
1020 Renens VD 1
Tel: (021) 635 08 65
Telefax: (021) 635 07 80

Librairie Payot S.A.

6, place Pépinet
CP 3212
1002 Lausanne
Tel: (021) 341 33 47
Telefax: (021) 341 33 45

Librairie Univers

6, rue de Candolle
1205 Genève
Tel: (022) 520 26 23
Telefax: (022) 529 73 18

Subscription Agency/Agence d'abonnements

Dynagress Marketing S.A.
38 avenue Vibert
1227 Carouge
Tel: (022) 308 07 89
Telefax: (022) 308 07 99

See also - Voir aussi

OECD Publications and Information Centre
August-Bebel-Allee 6
D-53175 Bonn (Germany)
Tel: (0228) 959 12 0
Telefax: (0228) 959 12 17

TAIWAN - FORMOSE

Good Faith Worldwide Int'l. Co. Ltd.
9th Floor, No. 118, Sec. 2
Chung Hsiao E. Road
Taipei
Tel: (02) 391 7546/391 7597
Telefax: (02) 394 9176

THAILAND - THAÏLANDE

Sukatt Sam Co. Ltd.
113, 115 Pong Nakhon Rd.
Opp. Wat Rajapoth
Bangkok 10200
Tel: (662) 225 9531/2
Telefax: (662) 222 5188

TURKEY - TURQUIE

Kultur Yayinlari Is-Turk Ltd. Sti.
Ataturk Bulvarı No. 191/Kat: 13
Kavaklıdere/Ankara
Tel: 428 11 40 Ext. 2458
Deimalehler Cad. No. 29
Besiktas/Istanbul
Tel: 260 71 88
Telefax: 434828

UNITED KINGDOM - ROYAUME-UNI

HMSO
Cen enquiries Tel: (071) 873 0011
Postal orders only
P.O. Box 276, London SW8 5DT
Personal Callers HMSO Bookshop
49 High Holborn, London WC1V 6HB
Tel: (071) 873 8203
Branches at: Belfast, Birmingham, Bristol, Edin-
burgh, Manchester

UNITED STATES - ÉTATS-UNIS

OECD Publications and Information Centre
2001 L Street N.W., Suite 700
Washington, D.C. 20036-4910
Tel: (202) 785 6323
Telefax: (202) 785 0350

VENEZUELA

Libreria del Este
Avenida Miranda 52, Apdo. 60337
Edificio Galpán
Caracas 106
Tel: 951 1705/951 2307/951 1297
Telegram: Libreria Caracas

Subscription to OECD periodicals may also be placed through main subscription agencies.

Les abonnements aux publications périodiques de l'OCDE peuvent être souscrits auprès des principales agences d'abonnement.

Orders and inquiries from countries where Distributors have not yet been appointed should be sent to OECD Publications Service, 2 rue André-Pascal, 75775 Paris Cedex 16, France.

Les commandes provenant de pays où l'OCDE n'a pas encore désigné de distributeurs peuvent être adressées à OCDE, Service des Publications, 2, rue André-Pascal, 75775 Paris Cedex 16, France.

1-1995

OECD PUBLICATIONS, 2 rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(91 95 02 3) ISBN 92-64-04355-1 - No. 47712 1995

MEASURING THE QUALITY OF SCHOOLS

How do we identify and measure how schools function? This publication examines this issue and asks how these functions – attribution of human resources, decision-making, the curriculum, internal organisation and instructional quality – can be expressed as comparable indicators. In its contribution to a better understanding and comparison of school operations across countries, this volume considers the basis for selecting and identifying indicators, their political relevance, empirical justification, strengths and weaknesses.



MESURER LA QUALITÉ DES ÉTABLISSEMENTS SCOLAIRES

Comment identifier et mesurer le fonctionnement de l'école ? Cet ouvrage tente de répondre à cette question en examinant comment l'attribution des ressources humaines, les prises de décision, les programmes d'enseignement, la qualité des méthodes d'organisation et d'instruction dans les établissements scolaires peuvent être exprimés sous forme d'indicateurs comparables. Afin de mieux saisir et comparer les systèmes scolaires des différents pays, cet ouvrage étudie comment sont choisis et déterminés ces indicateurs, leur utilité pour les pouvoirs publics, leur justification empirique, leurs avantages et leurs insuffisances.

(91 95 02 3) UDX
ISBN 92-64-04355-1



237